SCIENCE LIBRARY

18

APR 4 - 1932

Vol. XV, No. 2 PSYCHOLOGICAL REVIEW PUBLICATIONS APRIL, 1932

Journal of Experimental Psychology

EDITED BY

SAMUEL W. FERNBERGER, UNIVERSITY OF PENNSYLVANIA

HOWARD C. WARREN, PRINCETON UNIVERSITY (*Review*)
WALTER S. HUNTER, CLARK UNIVERSITY (*Index*)
HERBERT S. LANGFELD, PRINCETON UNIVERSITY (*Monographs*)
EDWARD S. ROBINSON, YALE UNIVERSITY (*Bulletin*)

CONTENTS

The Responses of Infants to Color: R. STAPLES 119

The Effect of Emotional Excitement upon Muscular Steadiness: W. N. KELLOGG 142

Effect of Pitch of Tone-Stimuli upon Body Resistance and Cardio-Vascular Phenomena: L. E. MISBACH 167

The Form of the Curve of Memorizing: G. C. WHITE 184

Anticipation of Correct Responses as a Source of Error in the Learning of Serial Responses: F. H. LUMLEY 195

Card Sorting as a Measure of Learning and Serial Action: M. A. TINKER, A. J. IMM AND C. A. SWANSON 206

Hypnotic Suggestion and the Conditioned Reflex: V. E. FISHER 212

A New Apparatus for Measuring Choice and Decision: M. F. HAUSMANN. 218

Apparatus for Measuring Activity: T. H. HOWELLS 226

PUBLISHED BI-MONTHLY
FOR THE AMERICAN PSYCHOLOGICAL ASSOCIATION
BY THE PSYCHOLOGICAL REVIEW COMPANY
PRINCE AND LEMON STS., LANCASTER, PA.
AND PRINCETON, N. J.

Entered as second-class matter, July 1, 1920, at the post-office at Lancaster, Pa.

PUBLICATIONS

OF THE

AMERICAN PSYCHOLOGICAL ASSOCIATION

EDITED BY

HOWARD C. WARREN, PRINCETON UNIVERSITY (*Review*)
S. W. FERNBERGER, UNIVERSITY OF PENNSYLVANIA (*J. Exper. Psych.*)
WALTER S. HUNTER, CLARK UNIVERSITY (*Index and Abstracts*)
HENRY T. MOORE, SKIDMORE COLLEGE (*J. Abn. and Soc. Psychol.*)
HERBERT S. LANGFELD, PRINCETON UNIVERSITY (*Monographs*)
EDWARD S. ROBINSON, YALE UNIVERSITY (*Bulletin*)

HERBERT S. LANGFELD, Business Editor

PSYCHOLOGICAL REVIEW

containing original contributions only, appears bi-monthly, January, March, May, July, September, and November, the six numbers comprising a volume of about 540 pages.

PSYCHOLOGICAL BULLETIN

containing critical reviews of books and articles, psychological news and notes, university notices, and announcements, appears monthly (10 numbers), the annual volume comprising about 720 pages. Special issues of the BULLETIN consist of general reviews of recent work in some department of psychology.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

containing original contributions of an experimental character, appears bi-monthly, February, April, June, August, October, and December, the six numbers comprising a volume of about 700 pages (from Jan. 1, 1932).

PSYCHOLOGICAL INDEX

is a compendious bibliography of books, monographs, and articles upon psychological and cognate topics that have appeared during the year. The INDEX is issued annually in June, and may be subscribed for in connection with the periodicals above, or purchased separately.

PSYCHOLOGICAL ABSTRACTS

appears monthly, the twelve numbers and an index supplement making a volume of about 700 pages. The journal is devoted to the publication of non-critical abstracts of the world's literature in psychology and closely related subjects.

PSYCHOLOGICAL MONOGRAPHS

consists of longer researches or treatises or collections of laboratory studies which it is important to publish promptly and as units. The price of single numbers varies according to their size. The MONOGRAPHS appear at irregular intervals and are gathered into volumes of about 500 pages.

JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY

appears quarterly, April, July, October, January, the four numbers comprising a volume of 448 pages. The journal contains original contributions in the field of abnormal and social psychology, reviews, notes and news.

ANNUAL SUBSCRIPTION RATES

Review: $5.50 (Foreign, $5.75). **Index:** $4.00 per volume.
Journal: $7.00 (Foreign, $7.25). **Monographs:** $6.00 per volume (Foreign, $6.30).
Bulletin: $6.00 (Foreign, $6.25). **Abstracts:** $6.00 (Foreign, $6.25).
Abnormal and Social: $5.00 (Foreign, $5.25). Single copies $1.50.
Current numbers: **Journal,** $1.25; **Review,** $1.00; **Abstracts,** 75c; **Bulletin,** 60c.

COMBINATION RATES (from Jan. 1, 1932)

Review and Bulletin: $10.00 (Foreign, $10.50).
Review and J. Exp.: $11.00 (Foreign, $11.50).
Bulletin and J. Exp.: $12.00 (Foreign, $12.50).
Review, Bulletin, and J. Exp.: $16.00 (Foreign, $16.75).
Review, Bulletin, J. Exp., and Index: $19.00 (Foreign, $19.75).

Subscriptions, orders, and business communications should be sent to the

PSYCHOLOGICAL REVIEW COMPANY

PRINCETON, NEW JERSEY

Journal of
Experimental Psychology

VOL. XV, No. 2 APRIL, 1932

THE RESPONSES OF INFANTS TO COLOR

BY RUTH STAPLES[1]

University of Nebraska

For many years the subjects of color vision and color preference in the infant and child have stimulated investigation. Since there has been great variety in the methods employed and in the brightness and saturation of the colored materials used it is not surprising that the results of the many experiments present some inconsistencies. The age at which color vision is possible has been set all the way from the fourth month to the end of the third year and there is considerable disagreement as to color preferences throughout childhood. The outstanding difficulties in investigating the problem appear to be as follows: (1) To secure a response which will be an adequate measure of color perception at an age when the infant is too young to use or to understand language. (2) To insure that the responses are to color and not to brightness. (3) To study a sufficiently large number of cases.

It is possible that in infancy color vision is still in an undeveloped stage. If such is the case a satisfactory measure of the perception of color might be of considerable theoretical importance. The present investigation has been undertaken for the purpose of determining the earliest age at which a response to color may be obtained, whether differences exist in the responses to the various colors, and at the same time

[1] From the Institute of Child Welfare, University of Minnesota. Acknowledgment is due to Dr. John E. Anderson and Dr. Florence L. Goodenough for many helpful suggestions and criticisms during the course of this investigation.

to overcome the difficulties listed above. Two hundred sixty-two different subjects under two years have been observed. The study has been conducted in three parts. Part I deals with infants under five months, Part II with the responses to color as compared with gray of infants from six to twenty-four months, Part III with the responses of infants of this same age to paired colors.

Part I

Subjects.—The subjects were 23 infants ranging in age from 69 to 143 days with an average age of 87 days. Two subjects were available long enough to repeat the tests after a lapse of a month, making a total of 25 series of tests.

Materials.—The 'Atlas' papers of the Munsell Color Co. were selected as a means of presenting the color to the infants, as these papers are available in standard degrees of brightness and saturation. The particular papers used are designated as red 5/10, yellow 5/7, green 5/7, blue 5/6 and neutral 5 ($N/5$). The numerator refers to the brightness or 'value' of the color and the denominator to saturation. The papers represent the most saturated colors possible in the uniform brightness of '5.'[2]

Procedure.—The method used consisted in presenting to the infant a colored and a gray disc of equal brightness and keeping track of the time he looked at the colored disc as compared with the gray one. The two were presented to him simultaneously on a gray background of the same brightness as the discs. A screen eight by sixteen inches was made of the Munsell 'Atlas' $N/5$ mounted on heavy cardboard, with a slit in the center eight inches long and one-eighth of an inch wide. The colors were mounted on cardboard discs three inches in diameter and the two discs to be studied were placed on the screen and fastened firmly six inches apart. The screen was attached at the lower edge to a flexible arm which was fastened to a crib. The infant to be observed was placed in the crib facing the screen. The experiment was carried out in a small darkened room, with only the subject and observer present. The screen was illuminated by a 200 watt daylight lamp under a sixteen inch opal diffusing globe (Graybar 99 glass), the globe being placed twenty-four inches from the screen and directly behind the infant. The screen was adjusted in each case so as to be

[2] The U. S. Bureau of Standards reports, after an examination of the reflection from the different Munsell papers that "careful measurements with a spectrophotometer verify in a remarkable degree the consistency of the Munsell values for the different hues." (Priest, I. G., Gibson, K. S., and McNicholas, H. J., *An Examination of the Munsell Color System.* Technological paper of the Bureau of Standards, 1920, No. 167.)

twelve inches from the baby's eyes. The experimenter stood behind the screen and observed the subject through the slit. Since the distance between the observer's eyes and the baby's eyes was only twelve inches the eye movements could be accurately noted. At the same time the observer was hidden from the infant's view. One stop watch for timing the duration of the test was placed on the crib easily seen through the slit. The experimenter, standing behind the screen first moved a bright object slowly back and forth across the screen to attract the child's attention, and held it for a second or two exactly in the center and then removed it. As soon as the infant's eyes were focused on one of the discs (color or gray), the time on the stop watch was recorded and the screen kept exposed for two minutes. During this period the number of seconds the baby looked at either the color or gray was recorded by two other stop watches held by the experimenter, one for color and one for gray, the appropriate watch being started when the eyes rested on the disc, and stopped when removed. After two minutes the screen was swung aside and the total time for each disc recorded. The baby was played with for a few minutes, then the position of the colored and the gray discs was reversed on the screen and the experiment repeated. The same procedure was carried out with each infant, each of the four colors being compared with gray. The order of presentation of the four colors was rotated with the successive subjects.

Results.—Table I presents a summary of the seconds accredited to each of the colors and to gray by the 25 subjects.

TABLE I[3]

COMPARISON OF ATTENTION GIVEN TO COLOR AND GRAY BY 25 INFANTS
(AV. AGE 87 DAYS)

Color Pairs Presented	Total Seconds	Av. Sec. for Each Subject	S.D.	r Color and Gray	Diff. in Means	σ Diff.	$\frac{D}{\sigma d}$
Red	602.2	24.09	22.90				
Gray	421.6	16.90	13.90	.75 ± .06	7.23	3.10	2.33
Yellow	678.2	27.13	27.91				
Gray	256.6	10.26	10.74	.68 ± .07	16.86	4.41	3.82
Green	504.4	20.17	25.38				
Gray	328.3	13.13	12.10	.64 ± .08	7.04	3.99	1.76
Blue	658.5	26.34	32.35				
Gray	294.6	11.78	10.30	.52 ± .10	14.57	5.85	2.49
Total color	2443.3	*24.43	*27.74				
Total gray	1301.1	*13.01	*12.05	*.60 ± .09	*11.42	*2.26	*5.05
Total color	2443.3	†97.73	†84.3				
Total gray	1301.1	†52.04	†36.3	†.83 ± .05	†45.69	†11.58	†3.94

* Averaged on basis of the number of tests (100, four for each subject).

† Averaged on basis of the number of subjects (25).

[3] Detailed tables with the responses of individual subjects are on file at the University of Minnesota Library.

A positive correlation was found to exist when the time the subject looked at each color was correlated with the time each subject looked at gray. This relationship may be explained in the following way: When a child's attention was attracted to the lighted screen, if he was sufficiently interested to look at the screen, both the colored disc and the nearby gray disc received a share of his attention. Consequently, in determining the standard error of the differences of the means, the formula for differences between two correlated measured was used.

It will be noted that each of the four colors was looked at for a greater number of seconds than the simultaneously presented and equally bright gray. When the totaled figures for colors and for gray are compared, the differences between the means are four and five times the $S.E._{diff.}$ for the two methods of computing total color, respectively. Consequently, these differences may be considered as significant. In the cases of the individual colors, the mean difference between yellow and gray is practically four times its standard error, and is hence a reliable one. With the other three colors, the differences, although within the limits of chance, are consistently in favor of the colors. It would appear then, from a study of these figures, that the sensation of color as differing from the sensation produced by a gray of the same form and brightness, may be experienced and responded to by infants as young as three months.

The rank of the colors, according to seconds looked at, is yellow, blue, red, green, but when the differences between these colors are subjected to the previously described statistical treatment they do not prove to be significant. This suggests that while the infants sensed color as differing from gray, there is little evidence that they responded selectively to the different colors. Sex differences were not apparent.

Check Experiment.—In order to determine (*a*) the accuracy with which the eye movements of the infants may have been recorded and (*b*) the effect of diffuse daylight instead of artificial lighting, the following check experiment was carried out. The situation of the main study was reproduced in

materials and equipment, save that light from a large north window was used to illuminate the screen. The tests were made only on clear sunny days in the month of June, between 2:30 and 3:30 P.M. The procedure of the original study was duplicated, excepting that two observers recorded the eye movements of each subject at the same time. Seven infants ranging in age from 79–147 days were observed. One subject was observed twice, making eight observations in all. One color and gray pair, only, was presented at each observation. The four colors were used successively, making two tests with each color.

(*a*) *Accuracy of recording eye movements.*—A correlation coefficient of .82 exists between the seconds credited to gray by the two observers in the sixteen two minute periods. When the seconds credited to color by the two observers are correlated, a coefficient of .92 is obtained. The difference between the seconds credited to color and gray by each observer in each of the test periods was estimated, and these differences correlated, the resulting coefficient being .87. In the case of the single observer the eyes were close to the slit and only twelve inches from the infant's eyes. When two persons looked through the slit at the same time it was impossible to be both close to the slit and directly in front of the child. The degree of relationship that existed even under these adverse conditions indicates that the method is a sufficiently reliable one to accept the figures as an adequate measure of the time that the infants looked at the colors.

(*b*) *Effect of diffuse daylight.*—Tables 2 and 3 present a comparison of the main study and the check experiment in regard to the length of time color and gray were looked at by the infants under the two conditions of lighting.

TABLE II

COMPARISON OF THE TIME CREDITED TO COLOR AND TO GRAY WITH ARTIFICIAL AND WITH DIFFUSE DAYLIGHT

	N	Av. Seconds for Color	Av. Seconds for Gray
Artificial daylight	25	24.4	13.0
Diffuse daylight (figures for the two observers averaged)	8	32.7	10.6

TABLE III

DIFFERENCES BETWEEN THE TIME CREDITED TO COLOR AND TO GRAY, WITH ARTIFICIAL AND DIFFUSE DAYLIGHT

	Diff. in Means	σ Diff.	$\frac{D}{\sigma d}$
Color	8.3	7.08	1.03
Gray	2.4	3.74	.64

It will be noted from the above tables that the responses of the infants to the color and to the gray were practically the same under both the artificial light and the diffuse daylight. The color was looked at very consistently more than the gray in both experiments. The differences between the means of the main and check experiments for color and gray respectively are not significant. Consequently it would appear that slight differences in brightness which may have existed between the colors and gray due to the particular lighting conditions employed did not materially affect the responses of the infants.

It may be well to emphasize that the extreme difficulty of securing and holding the attention of the infants of so young an age brings with it many chance factors. The great variability in the time of response given to the colors is doubtless explained in this way. But the infants were consistently more responsive to color than to gray in the case of each color and under both conditions of lighting, notwithstanding the difficulties involved in securing their attention. This fact indicates the presence of some degree of color vision at the age of three months.

PART II

The Responses to Color as Compared with Gray. Infants Between Five and One-Half and Twenty-Four Months of Age

Subjects.—One hundred eighteen different infants were studied, six being observed a second time after three or four months had elapsed, making a total of 124 tests. The subjects were tested while in attendance at one of the clinics

of the Minneapolis Infant Welfare Society. As these clinics are preventive in purpose, no sick infants are admitted. The subjects represent very closely a middle section of the population according to information obtained concerning paternal occupation.

Materials.—The 'Atlas' papers of the Munsell Color Company were used in Part II as in Part I, red 5/10, yellow 5/7, green 5/7, blue 5/6 and *N*/5. Charts of heavy neutral gray cardboard eighteen by twenty-four inches were prepared in the following manner: A sheet of 'Atlas' neutral '5,' size 9 × 14 in., was first mounted in the center of the cardboard to serve as a background for the colored discs. On this two discs of paper two and one-half inches in diameter were very carefully mounted, each three inches on either side of the exact center of the card. One was colored, the other gray. Both were of the same brightness as the gray background. There were four charts in all, one with red and gray discs, one with yellow and gray discs, one with green and gray discs, and the fourth with blue and gray discs.

Since most babies as soon as they are able to grasp reach for every attractive object near at hand, it was assumed that if the subjects in this study saw the colored discs as distinct from the equally bright gray background, they would try to secure them when the chart was presented. Consequently the following procedure was carried out.

Procedure.—The subject was brought into a small, well-lighted room and seated on a table. Directly behind this table was a large window. The baby was never in direct sunlight and neither did the direct sunlight fall on the chart. The experimenter stood in front of the baby and presented the charts to him one at a time, holding them fifteen inches from the child's body, and at an even angle, so that the colored and gray discs were equally distant from him and neither one favored in position. The experimenter presented the first chart to the child. If the child did not respond soon he was urged to 'get the ball,' or 'show me the ball.' As soon as the child responded the card was removed from his sight, and the type of response recorded.[4] This procedure took place four times for each card, with the colored circle

[4] A preliminary study showed the following types of response to occur:

(1) Responding to the color by pointing or grasping.

(2) Responding to the gray disc or to the neutral portion of the chart by pointing or grasping.

(3) Distractions (emotional disturbance, grabbing or slapping the card, disturbance by outside stimuli).

in the four positions of left of the gray circle, at its right, above it and below it. Then the child was played with for a few minutes, and the second chart, with a different color contrasted to gray, presented similarly. This procedure was repeated with the third and fourth charts, thus comparing all four colors with gray, sixteen presentations in all. The test never lasted more than about ten minutes after rapport had been established and all four charts were presented during one experimental period. The tests were always made during clear, sunny weather, between the hours of 10:00 A.M. and 2:00 P.M. The experimentation was carried out between March 28 and August 25, 1929. Infants under nine months who failed to pass a preliminary test to determine their ability to grasp were not included in the study. This introduced a slight selective factor in the youngest group of infants.[5]

Results

The subjects have been divided into five age groups. Age group I includes the children from six to eight months inclusive; group II, those from nine to eleven months; group III, those from twelve to fourteen months; group IV, those from fifteen to eighteen months and group V, eighteen to twenty-four months. The results are summarized in Table IV. The responses of the infants have been grouped into

TABLE IV

Responses to Colors when Compared with Gray.
Infants 5½–24 Months. $N = 124$

Colors Compared	Age Group	I (6–8 mos.)	II (9–11 mos.)	III (12–14 mos.)	IV (15–18 mos.)	V (19–24 mos.)	Totals
	Type of Response	$N=28$ Percent	$N=27$ Percent	$N=22$ Percent	$N=22$ Percent	$N=25$ Percent	$N=124$ Percent
Red.....	Reaching for color...	68 8	80.3	86.8	94.8	93.0	83.7
Gray.....	Other responses.....	31.2	19.7	13.2	5.2	7.0	16.3
Yellow.....	Reaching for color...	59.3	81.8	84.0	92.6	100.0	81.7
Gray.....	Other responses.....	40 7	18.2	16.0	7.4		18.3
Green.....	Reaching for color...	45.9	60.7	76.8	88.1	90.1	70.9
Gray.....	Other responses.....	54.1	39.3	23.2	11.9	8.9	29.1
Blue.....	Reaching for color...	49.5	70.1	75.4	94.7	92.2	74.9
Gray.....	Other responses.....	50.5	29.9	24.6	5.3	7.8	25.1

(Distractions not included in estimating percentages)

[5] Jones, M. C. (The Development of Early Behavior Patterns in Young Children, *Ped. Sem.*, 1926, 33), reports 75 percent of the infants studied by her to be able to reach and grasp by 173 days, and 100 percent at 269 days. Gesell, A. (*Mental Growth of the Pre-School Child*, 1925), rates reaching for a spoon 'B' for 6 months old children, 65–85 percent being able to secure it at this age, and 'C' for the 9 months level, 85–100 percent reaching accurately.

two main classes to facilitate comparison: (1) the definite pointing and reaching reactions to the color, and (2) all other responses which might be indicative of a failure to notice the color, e.g. responding to the gray disc, to the center of the card, or no response at all.

Two facts are evident from an examination of these tables: *First*, that there is a steady increase in the positive responses to each of the colors up to the fourth age group (fifteen to eighteen months), and a corresponding decrease in the responses that suggest possible failure to see the colors or insufficient interest in them to respond by reaching. *Second*, that there are consistent differences in the responses to the individual colors up to this same age level. Red is responded to the greatest number of times, yellow is a close second, blue third, and green fourth.

There are several factors which may have entered into the baby's response of reaching or pointing to the colored disc. First of all, it was necessary that he see the color as something different from its background and from its companion gray disc. Second, he must have been able to coördinate his eye and hand movements sufficiently well to reach the small disc accurately. If he happened to touch the center of the card, though his eyes were on the color, the response was rated as negative. Third, he must have been willing to reach repeatedly (sixteen times) for the 'ball,' even though he never actually secured it. The method of evaluating the responses thus represents an extremely rigid method of scoring. It would seem, after a careful study of the figures in these tables, that color becomes an increasingly effective stimulus between the ages of six and fourteen months. The consistent differences in the responses to the four colors, especially at the first three age levels, suggests either a better vision for red and yellow than for the blue and green, or a greater preference for these colors which would stimulate more accurate and a greater amount of reaching. It is difficult to explain why the younger babies reached so much more for a red disc, when it was presented on a gray ground, than they did for an equally bright and similarly presented green disc,

unless they saw it better. The older babies, it will be noted, reached for all the colors about equally well.

In order to determine whether or not interest was sustained throughout the test period equally well for the four colors, the data were analyzed on the basis of the first, second, third and fourth presentation of each chart. There was less reaching with each successive presentation of the chart for all four colors. The interest was maintained longer for red and yellow than for blue and green, the drop in the percentage of reaching from the first to the fourth presentation being least in the case of red, next with yellow, third with blue, and greatest in the case of green. This tendency was most marked in the case of the youngest subjects, and was evident only to a slight degree after the age of fourteen months.

Sex Differences.—With every color, and practically in every age group, the girls score a higher percentage of positive reaching responses than the boys, although the differences are not great (Fig. 1). The differences are larger between the sexes for the colors blue and green than for red and yellow. Both sexes respond more to red and yellow than to blue and green, as did the group as a whole. The girls appear to reach a high level of response to color earlier than the boys. This difference may be due to a more rapid development of color vision in girls, or provided equal ability to see the colors exists, may indicate a greater interest in color, and consequently a greater responsiveness. The generally accepted fact of greater acceleration in the development of females from birth to maturity offers a possible explanation.

No differences appear between the responses of the children of different social status groups.

Reliability.—In order to determine the amount of experimental error which might exist in the method of recording the responses of the infants, three observers took part in an experiment similar to the main study, and individually scored 12 subjects. Each chart was presented to each subject sixteen times, making a total of 192 judgments by the observers. The judgments of observers A and B showed an agreement of 93.2 percent, those of observers A and C,

95.9 percent, and those of B and C, 92.2 percent. The method of scoring the responses evidently was a very objective and reliable one.

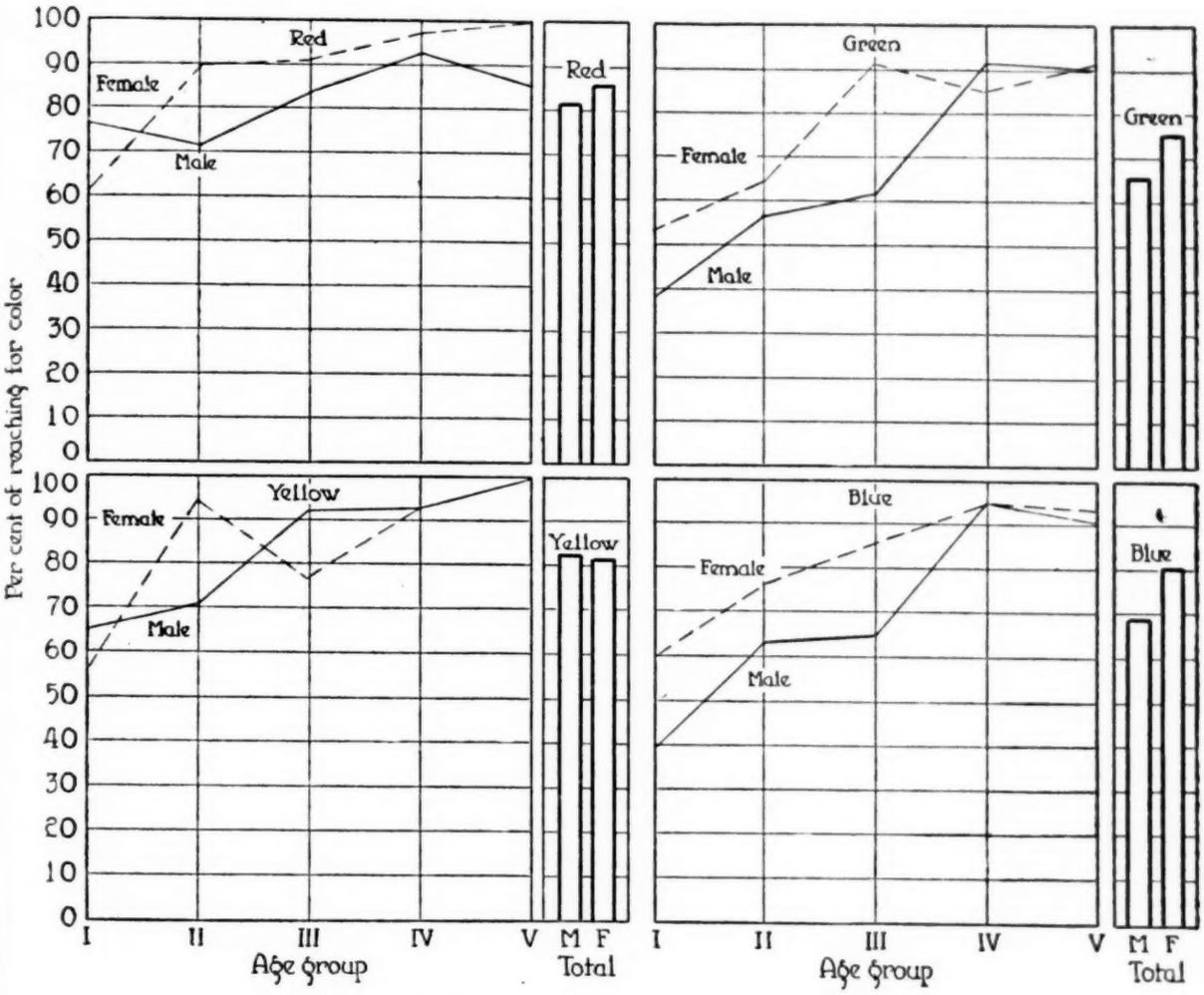


FIG. 1. Comparison of sexes in response to color.

Effect of Slight Difference in Brightness Between the Color and Gray.—Slight variations in the illumination may have occurred either in the individual experiments or between the conditions of standardization of the papers and the experimental conditions. Because of this fact, it was deemed advisable to determine whether slight differences in brightness between the color and the gray would in any way influence the child's response. Four charts were made by the Munsell Co., corresponding in general to the charts described as used in the main study. This time, however, the gray disc was very slightly brighter than the colored disc and the gray

background.[6] Twenty subjects were selected to duplicate in age and sex the main group of children. Two boys and two girls were in each of the five age groups. The average age of these infants was 13.2 months, as compared with 13.3 months for the main group of children. The procedure and conditions of the main experiment were reproduced. Table V presents the summary of the results with these twenty subjects, and compares the main and check experiments.

TABLE V

COMPARISON OF RESPONSES IN MAIN STUDY AND THOSE WHEN
(*A*) GRAY WAS BRIGHTER THAN THE COLOR, AND
(*B*) GRAY NEARER THE SUBJECT THAN THE COLOR

	Responses to Red Percent	Responses to Yellow Percent	Responses to Green Percent	Responses to Blue Percent	All Responses to Color Percent
Main Study	83.7	81.7	70.9	74.9	77.7
Check Experiment A (Gray brighter than color)	85.3	80.6	71.8	72.9	77.8
Check Experiment B (Gray nearer subject than color)	84.9	76.8	70.8	80.8	78.7

Evidently the slight difference in brightness had no material effect on the responses of the subjects, for with each color the total reaching is practically the same, whether the gray disc was brighter as in this check test, or of the same brightness as the background and color as in the main study. The reaching shows steady increase with age up to fourteen months as before, and the order of the colors is identical, red, yellow, blue and green. Thus we may assume with a fair degree of certainty that the responses of the subjects represented a true response to color, and not to possible slight differences in brightness between the color and the background.

Influence of Slight Changes in the Position of the Chart.— The charts were held so as to be parallel and at right angles to the edge of the table, but the infants naturally did not keep perfectly still. Consequently another check experiment was

[6] Munsell Gray No. 28, representing one-fourth of the distance between steps *N*/6 and *N*/5 of the black-white series.

undertaken to test any possible effect of slight differences in the distance between the colored and gray circles. Again, a group of twenty children were selected to match the main group in age and sex. Ten boys and ten girls were observed, four within each of the five age levels. The mean age of these subjects was 14.5 months, as compared with 13.3 months in the main study. The experimental materials and procedure were as before, and the conditions were duplicated, excepting that the charts were held each time so that the distance between the child and the colored disc was approximately three inches greater than the distance between the child and the gray. Table V shows the results of this second check experiment. The responses to the colors are very similar in both groups, although the positions of blue and yellow are interchanged. It seems probable from the above data, that any slight favoring of position, due to errors in experimental procedure or the changes in the position of the subjects had little, if any, effect on the results of the study.

Part III

Paired Comparison of Colors

Subjects.—One hundred twenty-one infants (60 girls and 61 boys) attending the same Infant Welfare Clinics as the subjects in Part II were observed. The age range was from $5\frac{1}{2}$ to 24 months, as before.

Materials.—Six charts similar to those used in Part II were used, made of heavy gray cardboard with the Munsell Neutral '5' serving as a background. Instead of having a gray and a colored disc, however, these charts each had two colored discs. The same Munsell atlas colors were used as in Part II, and each of the four colors was compared with every other color.

Procedure.—The room, lighting and position of the subject and experimenter were the same as in Part II. The first chart was presented to the child, and he was told to 'find the prettiest ball,' or if he were too young to understand this, the chart was merely held before him. In either case, the first color he reached for and touched or to which he pointed was considered as a response in favor of that color. The five other charts were then presented in a similar manner and one trial only was given for each chart. The order of presentation of the six charts was rotated with the successive

subjects, and the positions of left and right were alternated, to avoid a particularly favorable position for any one color.

If the child were emotionally disturbed, or responded merely by pushing the card away, the attempt to carry out the experiment was discontinued. Only well babies, who appeared to be normal, were studied. No children whose parents were known definitely to be feeble-minded were included. The colored papers used in all parts of this study were renewed frequently.

Comparative Study of Children and Adults.—Although considerable work has been done concerning the color preferences of school children and adults, in order to have more exact comparative data the experiment was performed upon three additional groups of subjects, one of 50 pre-school children, ranging in age from 2 years, 6 months, to 4 years, 10 months, with a mean age of 46 months, and one of 100 public school children from grades 3, 4 and 5, ranging in age from 7 to 12 years, with a mean age of 9 years. The third group was composed of 100 college students. The experimental conditions were practically the same as for the babies, and duplicate charts were used for the color comparison. Each subject was tested individually, being told to point to the one of the two colors on each chart which he preferred. The same care was taken to rotate the order of the chart presentation, and to alternate the position of the colors, as with the infants.

Results—Infants.—Table VI presents the summarized experimental data for the infants, the subjects being classified into the same five age groups as in Part II. Figures 2 and 3 represent the scale values for the colors compared. The method advocated by Thurstone [7] for showing the scale values for comparative judgments has been used here.

[7] Thurstone, L. L., An Experimental Study of Nationality Preferences. *J. General Psychol.*, 1928, 1, 405-425. The rank order of the preference is determined, and a table of sigma values prepared from the percentages credited to each color, the value being positive when the percentages are over 50 and negative when below. The formula:

$$(S_1 - S_2) = \frac{\sqrt{2}(\Sigma x_{1K} - x_{2K})}{N},$$

is then applied, S_1 and S_2 representing any two judgments where the scale separation is to be determined and x_{1K} and x_{2K} the difference in sigma values between any other judgment and S_1 and S_2 respectively.

TABLE VI

PERCENTAGE OF POSITIVE REACHING RESPONSES TO PAIRED COLORS. INFANTS 6–24 MONTHS

Age Group	I	II	III	IV	V	
	6–8 mos. $N = 26$	9–11 mos. $N = 26$	12–14 mos. $N = 23$	15–18 mos. $N = 23$	19–24 mos. $N = 23$	All Cases $N = 121$
	Percent of Responses	Percent of Responses	Percent of Responses	Percent of Responses	Percent of Responses	Percent of Responses
Red	64.0	73.1	65.2	52.2	52.2	61.7
Yellow	36.0	26.9	34.8	47.8	47.8	38.3
Red	83.3	65.4	60.9	59.1	43.5	62.7
Green	16.7	34.6	39.1	40.9	56.5	37.3
Red	72.0	69.2	68.2	59.1	73.9	68.6
Blue	28.0	30.8	31.8	40.9	26.1	31.4
Yellow	60.0	46.2	50.0	59.1	47.8	52.5
Green	40.0	53.8	50.0	40.9	52.2	47.5
Yellow	54.2	65.4	63.6	57.1	43.5	56.9
Blue	45.8	34.6	36.4	42.9	56.5	43.1
Green	45.8	46.1	34.8	36.4	34.8	39.8
Blue	54.2	53.9	65.2	63.6	65.2	60.2

Differences in Responses to Colors.—The most outstanding color preference of the infants is red (Table VI, Fig. 2). At each age level it is always preferred, whether compared with yellow, green or blue. This difference between the percentage of responses is 3.7 times the S.E.$_{diff.}$ when red is compared to yellow, 6.3 when compared with green, and 6.2 when compared with blue.[8]

Blue is definitely preferred to green, and this again is a reliable difference, being 3.2 times the S.E.$_{diff.}$. This preference is greater in the last three age groups than in the younger groups. Yellow is preferred to blue and to green quite consistently although these differences are smaller and less reliable. There is with exception of the interest in blue, a very definite tendency for the preference between the colors to become less marked as age increases. This tendency is well illustrated in Fig. 2, showing the scale values of the

[8] σ diff. $= \sqrt{\frac{P_1 \times Q_1}{N} + \frac{P_2 \times Q_2}{N}}$.

colors for the infants. With each successive age group there is less difference between the best and least liked colors.

Sex Differences.—In no case of a color pair is the difference between the responses of the sexes pronounced. There is a

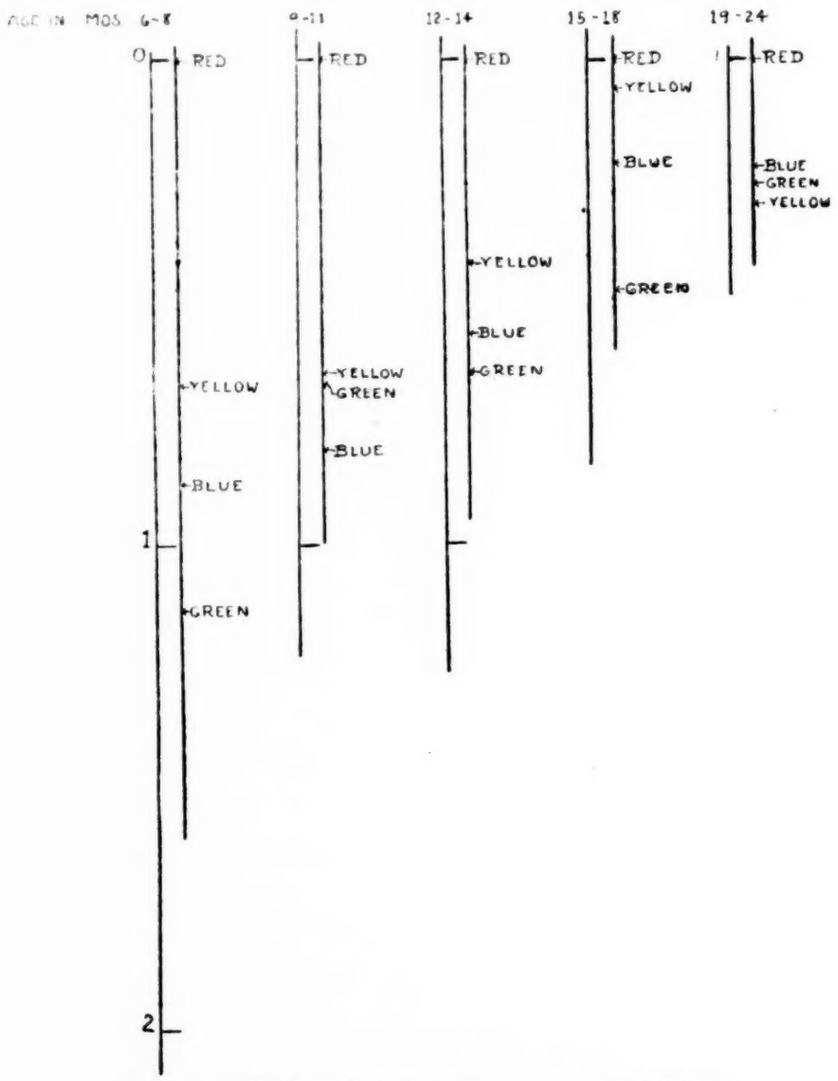


FIG. 2. Scale values for colors compared infants.

tendency for an earlier drop in the interest in yellow on the part of the girls. With the blue-green pair, the boys' interest is almost equally divided throughout the age range, but the girls have a greater percentage of responses for blue at each age level. This interest becomes stronger in the girls with increase in age, and it is the best liked color by nineteen months. The drop in the value of yellow for the girls is noticeable especially with the yellow-green and yellow-blue

pairs, and it is the least liked color by nineteen months. This lack of interest in yellow is not evidenced by the boys. As will be noted in Fig. 3, the yellow is the least liked color of the pre-school children, is less liked by the grade school

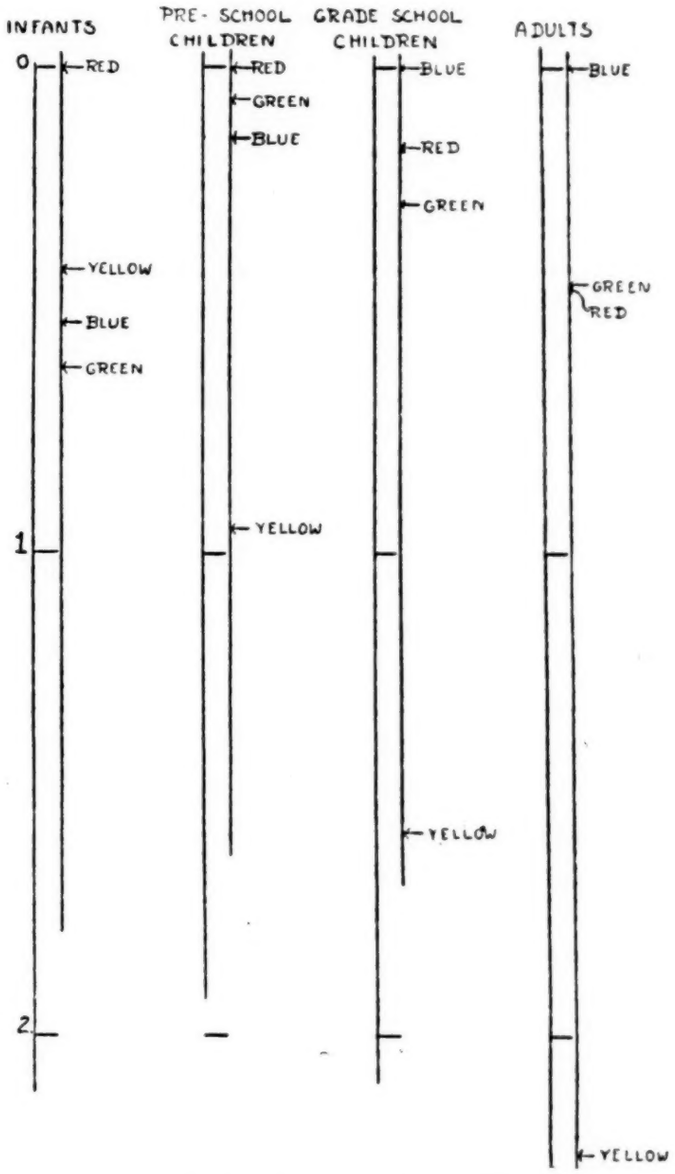


FIG. 3. Scale values for colors. All groups.

children, and still less liked by the adults, while blue is best liked by the grade school children, and still more strongly preferred by the adults. Thus these tendencies in regard to blue and yellow noticeable in the girl babies, may indicate a further progress along the color preference scale than is

shown by the boys of the same age. No consistent differences are shown when the responses of the upper socio-economic groups are compared with the lower.

Pre-School Children (Fig. 3).—In the case of the pre-school children, red, green and blue are preferred in the above order, but the differences between these colors are all small. Yellow is the least preferred color, and that this is a true lack of preference is shown when it is compared with red and blue, where the differences between the percentage of responses are 11.7 and 4.3 times the $S.E._{diff.}$ respectively. When yellow and green are paired, green is preferred, but the difference is not so large, and less reliable. The boys rate green a little higher than do the girls, while the girls rate yellow lower than the boys. No other sex differences appear.

School Children (Fig. 3).—In the case of the school children, blue leads, with red and green second and third. Yellow is again the least liked color, these children rating it considerably lower than did the pre-school children. When it is compared with red, green or blue, the differences between the percentage of responses are 20, 8 and 14 times the $S.E._{diff.}$ respectively. Blue is preferred when paired with red, this difference being 3.5 times the $S.E._{diff.}$. The differences between the blue-green and red-green pairs are unreliable. Yellow is rated again slightly lower by the girls than by the boys, and red slightly higher, but otherwise sex differences are not apparent.

Adults (Fig. 3).—The adults of both sexes rate yellow still lower than do the grade school children. When yellow is compared successively with red, green and blue, the difference in the percentage of responses is 22, 24 and 24 times the $S.E._{diff.}$ respectively. Both sexes prefer blue to red, when these colors are paired. With the men this difference is six times the $S.E._{diff.}$ but with the women only two times. The women rate green slightly higher than blue, but the men definitely prefer blue to green, the difference between the percentage of responses to these colors being eight times the $S.E._{diff.}$.

A careful comparison of the three older groups with the

babies reveals two marked changes in color preferences that are closely associated with age. *First*, the difference between the best and least liked colors decreases up to the age of two years, when the drop in the value of yellow brings about a steady increase between the greatest and least preferred colors up to the adult level. *Second*, there is distinct age relation to the colors which are preferred. There is a continuous, steady rise in the appreciation of blue, red is best liked until school age and then drops, and most marked of all is the drop in yellow, beginning at about eighteen months and steadily becoming more prominent.

Discussion and Conclusions

The results of this study suggest that color may be perceived, at least to a certain degree, by the age of three months.[9] Further, color becomes an increasingly effective stimulus until the end of the fourteenth month. Probably by twelve and at least by fifteen months, saturated red, yellow, blue and green are seen clearly. Between the ages of six and fifteen months, the different colors vary in their effectiveness in bringing about a response from the infants, the order being red, yellow, blue and green. This order is consistent throughout the experiment, appearing when the colors are paired with gray, in the paired comparison of colors, in the maintenance of the interest of the subjects and in the sex difference.

With the suggestion of a developing color sense, one immediately questions whether or not the infants in any way resemble the color blind. There are two possible resemblances when we consider the three months old subjects. *First*, these infants may have seen the colors merely as shades of gray as do totally color blind individuals, and their responses may have been to brightness differences existing between the

[9] Our information concerning the anatomical development of the retina indicates that color vision is quite possible at the end of the third month. Mann, J. C. (*The Development of the Human Eye*, Cambridge University Press, 1928), reports the cones of the retina to be present and well developed at this age, although the structural development of the central portion of the retina is not complete at least until the end of the fourth post-natal month.

colored discs and the background. However, if this were true, red should have attracted more attention than the other colors as the contrast between this disc and the background would have been the greatest. It will be remembered that red actually held third place. *Second*, there is a suggestion of red-green weakness in the preference order for the four colors, yellow, blue, red, green, but this order may have been merely the result of chance. We will next consider the infants at the age where the differences between the colors were greatest (6–12 mos.). Here the order of preference was red, yellow, blue, green. The check experiment where the gray disc was brighter than the color or background showed that slight differences in brightness had no material effect on the responses of the infant. This would discount an explanation of the subjects' reactions on the basis of total color blindness as suggested above. Considering possible partial color blindness, if the babies were "deuteranopes" and saw only yellow and blue in the spectrum why should they have preferred red to yellow when they were presented as a pair? If they were "protanopes" seeing the red end of the spectrum as black they might have been reacting to a brightness contrast in the case of the red disc, but it seems improbable that they would fail to grasp the best seen colors, blue and yellow less well than red. And it does not explain why yellow was reached for more than blue. Some of the recent studies made on color blind subjects, however, report that failure to recognize green is much more common than failure to recognize red.[10] One cannot say that the infants in this experiment possess the same 'green weakness' as do the color blinds. Since green is the color which presents the greatest difficulty in cases of color blindness and is also the color to which the infants are consistently the least responsive, it is possible that a fundamental relationship exists between the two which at present is not adequately explained.

[10] Sulina, A., Ueber den Farbensinn bei Kindern Nebst Bemerkungen über die Untersuchungsmethoden. *Russ. Opthal. J.*, 1923, **2**, 26. Collins, M. *Color Blindness*, New York, Harcourt, Brace & Co., 1925. Terman, S. W., A New Classification of the Red-Green Color Blind. *Am. Jour. Psychol.*, 1929, **41**, 237. Miles, W., One Hundred Cases of Color Blindness Detected with the Ishihara Test. *Jour. Gen. Psychol.*, 1929, **2**, 535.

It may be of interest to consider how satisfactorily the more well known color theories explain the behavior of our subjects. The explanation of color blindness offered by the Young-Helmholz theory (*i.e.* an absence or diminution of sensitiveness of one of the light-sensitive substances), suggests a possible solution. It could be assumed in the case of the infants that, provided these three substances exist, the red component is the most strongly developed, that the blue component is in a less well developed stage and that the green is weak. While this is theoretically possible, the explanation however, might be criticized as involving too many hypothetical assumptions to be very convincing. It is exceedingly difficult to explain our results according to the Hering theory, especially in regard to the positions of red and green. If the sensations of red and green are due to different processes in the same retinal substance, as this theory postulates, they should be sensed in equal proportions at the same stage of development. This is not the case in this particular investigation. Neither does the Ladd-Franklin theory explain the responses of the infants. There was a slight indication that the youngest infants saw blue and yellow better than red and green, but this difference may have been one of chance. Why is red such an effective stimulus at six months while green lags behind?

In regard to the drop in the interest in yellow between two years and the adult level the particular color may have been a contributing factor. It was darker than spectral yellow and may have been peculiar from the adult standpoint in so far as it did not conform to their idea of a good yellow. It is interesting to note that the older children and adults often referred to it as 'tan,' while the pre-school children called it 'yellow.' However, in other studies of color preferences where a spectral yellow has been used this same drop has been observed.[11]

Inasmuch as the findings in regard to the effectiveness of

[11] Holden, W. A., and Bosse, K. K., Order of Development of Color Perception and Color Preference in the Child. *Arch. of Opthal.*, 1900, **29**, 261. Katz, S. A., and Breed, F. S., The Color Preferences of Children. *J. Appl. Psychol.*, 1922, **6**, 255. Michaels, G. M., Color Preference According to Age. *Am. J. Psychol.*, 1924, **35**, 79.

the various colors as stimuli are markedly consistent from child to child, from age to age and in the different groups of children observed in the various parts of this study, they cannot be interpreted as a result of chance. The youth of the subjects indicates that the variation in response is not due to training or experience. Since the most marked differences in response occur between six and twelve months, and since beyond this age, group trends begin to disappear and individual differences become more pronounced, some phase of the development of color vision appears to be in progress during the first year of life. That some fundamental factors are involved which would bear further investigation appear evident as a result of this demonstration of an extremely consistent age relationship in the reactions of infants to different colors.

Summary

1. Two hundred sixty-two infants ranging in age from 69 days to 24 months were observed to determine any differences in their responses to equally bright red, yellow, green, blue and gray. Fifty pre-school children, one hundred grade children and one hundred adults were observed for comparison.

2. The colors were presented on a gray background (*a*) with gray and (*b*) with each color paired with every other color. With the infants too young to respond by grasping, the time they looked at the colors and gray was recorded. With the infants able to reach, the colors were presented for grasping.

3. The sensation of color as distinct from that of gray may be experienced by infants as early as the end of the third month of life. At this age no significant differences in responses to the four colors were observed.

4. The four colors were clearly seen probably by twelve and unquestionably by fifteen months and at this age were experienced as distinctly different sensations.

5. Infants between six and twenty-four months are unequally responsive to the different colors. Red is the most effective stimulus, followed by yellow, blue, and then green.

These differences are most marked between six and twelve months.

6. There is a sex difference in responsiveness to color in favor of the girls which is greatest in the case of blue and green. The girls also show slightly more mature color preferences than do the boys.

(Manuscript received March 9, 1931)

THE EFFECT OF EMOTIONAL EXCITEMENT UPON MUSCULAR STEADINESS

BY W. N. KELLOGG
Indiana University

It is commonly supposed with reference to the emotions that they not infrequently express themselves through bodily trembling, unsteadiness and loss of motor control. Our everyday language is in fact rich in such statements as "he was so frightened his teeth chattered," "I felt my knees begin to shake," and "she actually trembled with anger." Undoubtedly the popular conception of 'nervousness' in temperamental make-up is related in some degree to the fact that certain individuals seem to be more subject to this type of activity than others—especially when in exciting, startling or disturbing situations. If then such reactions have received sufficient notice from the layman to contribute to our style and manner of speech, one may rightfully inquire why they have not been more thoroughly studied by the experimental psychologist. It appears indeed as if such obvious responses should afford a simpler and more direct means of examining emotional behavior than those elaborate procedures which involve the measurement of systolic blood pressures, inspiration-expiration ratios, cardiographic records, electrical skin reflexes and visceral contractions.

Performance in *voluntary* motor tasks as for example in reaction time and speed of tapping, has been recorded following emotional stimulation by Marston (**10**), Landis, Gullette and Jacobsen (**9**) and others, but these tests it seems to us fall in a category quite distinct from the measurement of incipient tremor movements. The latter type of behavior on the other hand which has been the subject of a variety of investigations appears nevertheless to have been little studied in its emotional aspects. Jastrow's use of the automatograph (**6**) probably represents one of the earliest

attempts to measure involuntary muscular responses to affective stimuli although the mildness of the situations he employed precludes for the most part their being regarded as genuinely emotional.[1] Later experiments by Fearing (**3**), Skaggs (**14**) and Laird (**7**) are more closely related to the present study. In many particulars, however, even these are directed towards different objectives so that only their relevant features will be reviewed here.

The first investigator using a Miles Ataxiameter (**11**) recorded individually the body sway of 2 groups of subjects over a series of 2-minute periods. The control group had simply to stand as steadily as possible while the experimental group had not only to remain steady, but was required in addition to count and later report upon a series of irregular taps which occurred during the test intervals. Curiously enough, Fearing found that the experimental group was reliably *steadier* than the control group even though it had a more complex *aufgabe* than the control group. The explanation of this fact appears from the introspections of the members of both groups which show that those subjects who counted taps (although their attention was distracted from the primary task of remaining steady) were less likely in the author's words to be 'panicky' than the subjects who were not so occupied. Distraction appears consequently to have a facilitating effect upon the steadiness of body sway instead of an inhibiting one.

Although little is said in Fearing's paper regarding the emotional influence of the distracting stimulus, it seems probable that such influence at the most was of an extremely low order. Skaggs (**14**), however, in an investigation of a different nature, presented stimuli which were purposely of a startling and exciting type. One of these, the 'startle' stimulus, was produced by a concealed automobile horn which was blown suddenly during the progress of the experiment. The other, calculated to arouse 'excited expectancy,' was an electric shock of which the subject had been previously

[1] For a general treatment of the literature on involuntary movements the reader is referred to Titchener (**15**, p. 158 ff.) and Whipple (**17**, pp. 155–60).

forewarned. Involuntary movements of the hand and arm were recorded with a Whipple Steadiness Tester (**17**, p. 157). A series of 15 sec. test periods before and after the stimulus gave steadiness counts from which it was possible to determine the effect of the stimulus. Although the respiration rate which was measured under the same conditions showed many of the subjects to be considerably affected, no characteristic changes in steadiness were observed. Skaggs concludes therefore that his results are chiefly negative in this respect.

The work of Laird (**7**) while less pretentious than that of Fearing and Skaggs offers a record of various motor performances *during* emotional upset. Laird had as subjects 8 pledges of a Greek letter fraternity whose scores on a Whipple Tester, and in 2 additional motor tasks of a *voluntary* nature, were measured while the subjects were (*a*) encouraged by their future fraternity brothers, and (*b*) heckled or 'razzed' by them. The steadiness records which covered only a 30 sec. period were considerably poorer (*i.e.* the contacts increased) during the 'razzing.' Laird's results may be criticized, however, on the grounds that it is impossible to tell from the procedure he employed whether the change in steadiness is a function of the distracting nature of the stimulation, of its emotional influence, or of both these factors.

Problem

The present study was undertaken therefore in the face of this brief and somewhat confusing history in the attempt to ascertain whether a more effective technique of arousing the emotions might not demonstrate after all significant steadiness changes *following* shock stimuli, which could be used to measure the degree of disorganization of the *S*s. It was further desired to obtain curves of the change in steadiness to indicate if possible the gradually decreasing effect of the disturbance as the subject returned to normalcy.

Experimental Conditions

a. Apparatus.—A plate-and-stylus tester of the Whipple type was selected to measure steadiness, since this device despite some disadvantages is probably the simplest one available. Slight modifications of the standard apparatus supplied by

Stoelting were made as follows: (1) a sliding metal cover was built in front of the contact plate to permit the exposure of only one hole at a time. This precluded the accidental insertion of the stylus into the wrong hole. (2) A 6-volt electric lamp was connected through a relay in the plate-stylus circuit so that it flashed whenever contacts were made. This light, mounted almost directly in *S*'s line of vision upon the upper edge of the contact plate, informed him when the stylus touched it and prevented him from falling unintentionally into the error of resting the stylus upon the circumference of the hole and so obtaining a spuriously high steadiness score. (3) The apparatus was mounted upon a small shelf-like platform on the *S*'s side of a solid partition in a windowless and partially sound-proof research room. The distance of the steadiness tester from the floor was equivalent to the shoulder-height of the average subject when seated in a special stimulus chair (*vide infra*). (4) The contact plate could be adjusted at various distances from the chair (which was fixed) to compensate for differences in the extended arm-length of individual subjects. (5) By means of a pivotal attachment the plate could also be tilted in either direction from the vertical so that *Ss* whose sitting height was unusually short or unusually tall could nevertheless extend their arms towards the tester at an angle perpendicular to the plate itself. No adjustments of more than 5° from the vertical were ever necessary during the course of the experiment, however.

Communication with the subjects except during the administration of the emotional stimuli was entirely by means of red and white signal lights arranged conveniently upon the partition a short distance above the contact plate. All accessory apparatus consisting of batteries, relays, stop-watches, an electric counter for recording contacts, and switches for operating signal lights were on the opposite side of the dividing partition where they were controlled by an assistant. Breathing and graphic time lines obtained simultaneously with the steadiness counts during parts of the experiment, were recorded with a kymograph and seconds pendulum which were also located in the assistant's section. The air tube from a Sumner pneumograph was run through the partition for this purpose.

b. Subjects and procedure.—Forty-two persons, 8 of whom were women, served as subjects. These were divided into 5 groups, 4 composed of 8 *Ss* each, while the fifth was made up of the remaining 10. The groups may be regarded as approximately equivalent in respect to the demands of the present experiment. Among the 4 units of 8 *Ss* each, the division of the sexes was the same. The age factor was also balanced as well as the 'degree of sophistication' as determined by the percentage of upper and lower classmen within each group. The members of the first 4 groups were tested separately with one of 4 different emotional stimuli. The fifth group of 10 *Ss* which was made up entirely of men, was used in a later variation of the regular procedure.

All the *Ss* were students from classes in psychology and the majority of them were freshmen. None had any knowledge whatsoever of the real purpose of the experiment before it was undertaken and each was required to sign a pledge of secrecy at its completion. The ostensible purpose, as far as the *Ss* themselves were concerned, was the measurement of steadiness over an interval of time. To this end the correct sitting posture, method of holding the stylus and signal system were carefully explained both by written and verbal instructions and were also demonstrated. The apparatus was separately adjusted to the dimensions of each *S* who was given practice trials until he was thoroughly familiar with the simple *modus operandi*.

Preliminary tests on about a dozen persons whose results are not treated here, had indicated that fatigue could be pretty well eliminated throughout a series of

steadiness records if the duration of each test period was not longer than 12 sec. with a 20 sec. interval between tests. An experimental series consisted of 12 such periods with their respective intermissions. When resting between successive tests the *S* was permitted either to place his arm upon his lap, or to move it if he wished to stimulate circulation.

The elimination of contacts which might be made while the stylus was being inserted or withdrawn was accomplished by not throwing the electric counter into the plate-stylus circuit until 2 seconds after the stylus had been properly placed by the subject. Correspondingly the assistant released the counter from the circuit *before* giving the light signal to withdraw the stylus. The actual recording of contacts therefore took place only during a 10 sec. period which was free from voluntary movements on the part of *S*.

An experimental sitting for any given subject consisted of 2 series of 12 tests each, one series being separated from the other by a rest interval of about 10 minutes. For 32 of the 42 *S*s, one of these series served as a 'control' in which no extraneous stimuli were given. After the fourth contact period of the other or 'stimulus' series, however, a surprise emotional stimulus was introduced. In order to balance practice or fatigue effects from one series to the other, the control tests were given first in the case of half the subjects, and last for the remaining half.

At the conclusion of the sitting, the *S* was taken to an adjoining room, the actual significance of the experiment was discussed with him, and a full introspective report was obtained. The entire procedure occupied on the average about one hour.

Emotional Situations

The shock stimuli used in this investigation were purposely selected from among those which have been proven in previous studies to be capable of inducing clear-cut emotional states. Four separate emotional situations were employed but no single individual was subjected to more than one of these.

a. Chair stimulus.—The first, hereafter called the 'chair stimulus,' is an adaptation of the fear situation employed successfully by H. T. Moore (**12**) and Blatz (**1**). A strong wooden laboratory chair whose two rear legs had been removed, was mounted upon an 8 inch platform to which the front legs were securely hinged at their bottom ends. The back of the chair seat was supported from beneath by a solid hinged brace that could be jerked out of position by a rope which ran through a series of pulleys to the assistant's side of the partition. When the hinged brace was removed the chair fell backwards through an angle of 60°. The complete operating mechanism, including the rope, was effectively concealed from *S* by a screen of black cloth about the lower part of the chair and by the platform to which the chair was fastened.

Although the chair was used by all subjects who made contacts and so formed an integral part of the experimental equipment, it was only 'dropped' with those *S*s who were specifically to be stimulated by this means. In such cases it was caught at the lower extremity of its fall by *E* who returned it to its original position. *E* therefore stood behind *S* who was aware of his presence but could not see him. To render this factor constant and to minimize its importance as much as possible, *E* was continually present throughout the stimulus and control series of every subject. He was, however, out of sight of *S*, except when the emotional stimuli were given, and he spoke only when absolutely necessary. *E*'s presence, consequently, must be considered an essential feature of each of the total situations.[2]

b. Snake stimulus.—The second stimulus, later referred to as the 'snake stimulus' has indeed been commonly employed in experiments of this sort and has frequently been reported to be very effective.[3] It was administered under the present technique by *E* who suddenly appeared beside *S* holding a five-foot chicken snake (*Coluber quadrivittatus*) in coiled position. Without warning, the animal was placed on *S*'s lap, where it was left for a period of 10 sec. before being removed.

c. Mouse stimulus.—The third, or 'mouse stimulus' is a modification of that originally used by Landis and Gullette (**8**) who required their subjects to decapitate a live albino rat. In the current procedure albino mice were substituted for rats, and *S* was made to cut the animal in two at its mid-section, instead of at the neck. In presenting this stimulus *E* stepped from behind S's chair and handed *S* the mouse securely wired by each leg to a 6″ by 10″ hardwood slicing board. *E* also gave *S* a dull 12-inch butcher knife. The following directions were repeated by *E*: "Here is a live mouse, and here is a butcher knife. I want you to slice the

[2] A better procedure would have been to eliminate any association between *E* and *S* during the experimentation but the application of some of the emotional stimuli would then have been impossible, as will presently appear. Such a complexity is illustrative of the difficulty of experimenting upon the emotions when the particular reactions of different *S*s cannot be accurately predicted.

[3] Notably by Brunswick (**2**).

mouse in two. Do not chop. Slice." Two of the subjects who refused pointblank to perform this task, were required to watch *E* do it.

d. Brains stimulus.—The fourth and last of these situations, here designated the 'brains stimulus' is a combination of H. T. Moore's technique of making the subject handle animal brains (**12**), with a strong electric shock. To the left of *S*'s chair as he seated himself before the steadiness tester was a small table on which rested a crockery jar 8″ in diameter by 12″ high with a wooden cover. Near the jar was a wooden pail to which had been specially fitted a hinged sheet-iron cover with a brass handle by means of which the cover could be readily lifted.[4] Inside the jar was concealed the hemisphere of fresh calf's brain, designedly chosen with as much congealed blood on it as possible. The brain section rested in about 3 inches of water. Inside the pail, and also invisible to *S* because of the closed cover was about 8 inches of water. Two hidden wires led from an inductorium on the assistant's side of the partition, to the pail, where they made contact respectively with the water in the pail and with its hinged metal top. The water and the cover consequently served as electrodes which if connected would complete the inductorium circuit.

In stimulating the subject with this equipment *E* stepped to *S*'s side, took the lid from the crockery jar and said: "With your left hand I want you to remove the brains which are in this jar and put them in the pail. Be sure to set them in the pail slowly for there is water in it which might splash." Since the brains were wet they moistened the subject's hand as soon as he touched them. It was necessary for *S* also to raise the metal cover of the pail with his right hand in transferring the brains with his left. As soon as the left hand touched the water in the pail therefore, a direct connection was made through *S*'s body to his right hand which held the metal cover of the pail. The inductorium circuit was thus completed, producing a vicious electric shock which was experienced chiefly in the right and left hands.

[4] Nothing was said about these articles at the start of the experiment which as far as the subject was concerned were incidental fixtures in the situation.

Results

a. The effectiveness of the stimuli.—Since the introspections of the subjects constitute the ultimate criterion for determining the type and degree of emotion induced, a question may rightfully be raised regarding their probable validity. Certainly the protocols of the inexperienced *Ss* used in this study cannot be considered as accurate as those of more advanced psychologists.[5] It is possible furthermore that the *Ss* in general approached the experiment in a more or less apprehensive condition which may have influenced their statements as well as their performance. We think this is rather unlikely, however, since *not one* of the reports indicates such a state and since the meticulous precautions taken to guard the true nature of the experiment eliminated any direct reason for suspicion.[6] The present technique of using each *S* for only one stimulus—a method at variance with that of most other investigations of emotional responses—was probably an important factor in this connection.

In support of the protocols, moreover, it should be noted that the subjects were not required to observe elusive thought processes or sensations, but only distinct and relatively prolonged affective states which should have been common enough in their own experience. Carefully directed questions were used to obtain all the facts from them and a list of possible emotions was presented as an aid. In addition they were cautioned to report nothing which was not clearly experienced. *E* is of the opinion therefore that in all but two instances he was able to win the confidence of the *Ss* who

[5] From one point of view, however, their naïvité may be looked upon as an advantage, since they were probably more definitely disturbed by the shock stimuli than would have been the case had they been more psychologically 'test-wise.' There was little evidence of the superior or cynical attitude which some investigators have found in subjects used in experiments of this type.

[6] Brunswick has aptly reviewed the chief aspects in which the introspections may be suspected of distortion as follows (2, p. 64): "They may not always be veracious, not that is, through any conscious effort of the reactor to fake or deceive. But he may be influenced by a certain amount of pride to cover up from himself as well as from the experimenter, especially his fear reactions, or at least to minimize them. . . . On the other hand the reactor may be influenced to exaggerate somewhat the reactions experienced, through a desire to please the experimenter with something positive."

coöperated sincerely in their introspective statements. The emotions tabulated for the 2 unsuccessful cases were those which *E* and the assistant judged were experienced even though they were not directly reported. With these minor exceptions therefore, we think the introspections can be relied upon.

Perhaps the most striking fact which appears from an examination of the protocols is the large variety of emotional states reported. This is corroborative in general of the findings of similar recent studies. Many of the *S*s it is true experienced nothing more than the characteristic 'surprise' or 'startle' which was recognized by all subjects regardless of the stimulus situation. Beyond this, however, the majority reported from one emotion to a complex consisting of several among which a particular feeling usually predominated. Following is a list of the outstanding terms used to describe the various subjective responses: amusement, anger, anxiety, annoyance, clammy unpleasantness, creepy feeling, determination to do better, disgust, distraction, disturbance, embarrassment, excitement, fear, horror, nausea, pity, repugnance, repulsion, resentment, shakiness, shock, shrinking feeling, startle, surprise, suspicion, unpleasantness, wonder, worry. [7]

There can be little question that severe emotional upset was produced in the cases of some of the *S*s, especially when the miscellaneous behavioral evidence which was frequently presented is considered. One male subject, in fact, who refused to 'slice the mouse in two' and who reported horror and nausea upon observing *E* do it, loosened his collar, perspired and slumped in his chair in a near-fainting condition. There was some doubt for the moment whether he would be able to continue the experiment, and although he did with evident difficulty complete the series he had to be given a drink and a rest upon its completion. A female *S* in the same situation, reporting pity and disgust, told *E* later that she had been unable to eat till the day following the experiment. Another female subject who experienced 'fear ap-

[7] It should be remembered that several descriptive terms were often employed in the introspections of any single *S*, and further that the same *S*s occasionally felt different emotions *at different times* following the stimulus.

proaching horror,' screamed 'Oh no!' when the snake was placed upon her lap, recoiled from it, put her hands over her face, and shuddered several times during the later rest intervals between contact periods.

It is by no means apparent from our results, however, that the women as a class were markedly more affected than the men. In fact, one of the female Ss when presented with the snake, although somewhat 'surprised' carefully examined it 'to find out if it was poisonous.' In this connection, the relative ineffectiveness of the snake stimulus to induce severe disturbances, was a revelation to the writer. The reports following this stimulus were usually of an extremely lukewarm variety, one subject even experiencing 'mild amusement.' Doubtless the general reception which the snake stimulus received is to be accounted for in part by the fact that a majority of students in this middle western university reside on farms, and so are more intimately acquainted with reptiles than an urban group would be.

b. The reliability of the steadiness records.—The first point to be considered with reference to the steadiness records involves their consistency. In the attempt to get an adequate measure of this factor, the reliability coefficients in Table I

TABLE I

RELIABILITY COEFFICIENTS OF STEADINESS RECORDS

	Stimulus Series	Control Series	Both Series
r	.947	.983	.976
PE_r	.006	.004	.004
N	42	32	74

were computed. These were obtained by correlating the total contact scores of the 6 even numbered periods with those of the 6 odd numbered periods, and correcting the resulting Pearson r by the Spearman-Brown prophecy formula.[8] Despite the relatively few cases in the stimulus and control series when these are separately considered, it will be noted that the rs are all exceptionally large.[9] It may

[8] *Cf. e.g.* Garrett (4, p. 269).

[9] The N for the control series is 10 less than that for the stimulus series since control records were not obtained for the entire 42 Ss.

be concluded therefore that under the technique of the present experiment, the steadiness measures possess a high degree of consistency.

c. Steadiness curves for different stimuli.—The comparison of the steadiness of the subjects during the various stimulus series with their steadiness during the control series is shown graphically in Fig. 1. Here the solid lines are made by

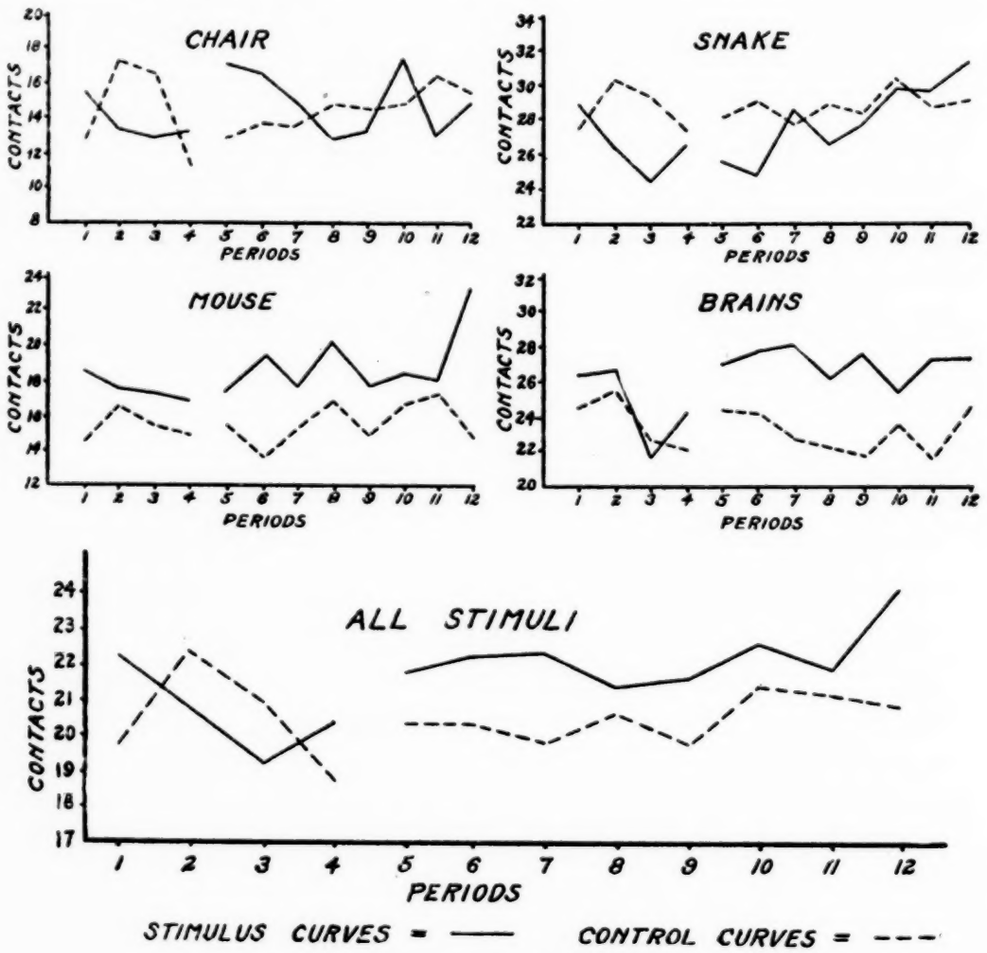


FIG. 1. Steadiness curves for the different stimuli.

plotting successive steadiness counts for the twelve 10-sec. periods when the emotional stimuli were introduced, and the broken lines represent the corresponding control records. A space is left in all curves between periods 4 and 5 to indicate the position in the stimulus series at which the emotional stimulus was presented.

At first glance the records of the 4 variously stimulated groups of 8 subjects each (upper part of Fig. 1) appear

somewhat perplexing. Thus although steadiness as shown in the stimulus curves sharply *decreases* (i.e. the number of contacts per 10 sec. increases and the curves therefore rise) following the mouse and brains stimuli, and to a lesser extent after the chair drop, there is no such abrupt change in the case of the snake stimulus. These differences can be accounted for on the one hand as the result of different degrees of effectiveness among the 4 stimuli. They can also be explained, however, by the fact that the subjects, although presumably 'equivalent' from one group to the next, may actually not have been so. The general tendency of the curves to rise toward the later periods is *prima facie* evidence that despite preliminary trials to determine the most desirable procedure fatigue was not entirely eliminated throughout the long series employed.

That the irregular fluctuations are chiefly the result of the few cases is demonstrated by the smoother curve which is produced when the results for the 4 smaller groups are pooled (lower part of Fig. 1). As might be inferred from the group curves, the combined tracing shows a small but distinct decrease in average steadiness following emotional stimulation. It is to be remarked, however, that throughout the approximately 4 minutes which elapsed after the stimuli had been presented (periods 5 to 12) the stimulus curve never drops to the level of the control curve although it tends to approach this level. The average disturbance in muscular control had not been entirely overcome therefore by the time the test series was concluded.

Although the predominant tendency as indicated by the pooled results is hence a decrease in steadiness for the 32 Ss taken as a whole, this can certainly not be taken to mean that the steadiness of every subject decreases following a surprise stimulus. A review of the records of the individual Ss shows that some in each of the 4 groups actually become *more steady*. Changes in this direction under the conditions of the present experiment, however, are exceptions rather than the rule and no infallible conclusions can be drawn from them. Nevertheless the fact that some persons *increase* in

steadiness is significant, particularly in the light of Fearing's earlier findings (3).

A relevant point in this connection is that many of the subjects who experienced no pronounced emotional state had a sudden insight into the true purpose of the experiment, accompanied by a determination not to permit their contacts to increase, soon after the stimulus was presented. The influence of motivation in cases of this sort is hence distinctly against an increase in contacts. If the *S* is not too seriously disturbed, therefore, he is likely to bring added effort to the task of trying to remain steady. This finding may possibly explain why some *S*s are more steady after stimulation than before, although it is conceivable, if the steadiness measured is entirely involuntary that an increase in effort (which supposedly is accompanied by greater muscular tension) may decrease the steadiness instead of increasing it. The complicating motivational factor must at any rate be responsible to some extent for the failure to obtain more clear-cut steadiness changes following the stimulation.

d. Shifts in variability.—If then there is not a characteristic change which is common to all subjects in the *number* of contacts, may there not be a uniform change in the *variability* of the contacts following stimulation? To answer this question each *S*'s results both for the control and for the stimulus series were divided into thirds, and the range within which the contacts varied throughout the periods making up each third was determined.[10]

The figures presented in Table II give the average ranges for the separate groups of *S*s in the different thirds of the stimulus and control series. The average range for all 32 subjects (as shown in the two lower lines of this table) increases during the stimulus series in periods 5–8 although the increase is no longer pronounced for periods 9–12. Inspection of the ranges for the separate stimuli discloses no uniformity in this regard, the range figures differing in a

[10] The separate division into thirds of control and stimulus contact periods for any given *S* allows only 4 measures in each third. Since 4 measures hardly warrant a precise computation of dispersion (*e.g.* the S.D., P.E., or A.D.) the range was here employed.

TABLE II

SHOWING AVERAGE RANGES IN NUMBER OF CONTACTS FOR CONTROL (C) AND STIMULUS (S) SERIES

Stimulus		Periods		
		1–4	5–8	9–12
Chair	C.....	9.50	7.63	8.38
	S.....	7.75	7.50	8.75
Snake	C.....	9.88	6.75	6.88
	S.....	10.38	10.00	8.00
Mouse	C.....	8.00	8.50	7.38
	S.....	8.00	12.25	8.75
Brains	C.....	8.25	8.38	6.75
	S.....	8.38	9.50	7.00
All stimuli	C.....	8.91	7.82	7.35
	S.....	8.63	9.81	8.13

manner analogous to the variation among the group curves of Fig. 1. The curves and the computations agree, however, in indicating a larger change following the mouse and brains stimuli than those which occur after the other two.

e. Affected vs. unaffected subjects.—In the attempt to throw further light upon the question of typical variations, the steadiness records were classified according to the kind of emotional state reported, without regard to the different stimulus situations. The results from this separation were also negative, showing no characteristic curves even for the so-called pleasant as opposed to the unpleasant emotions. (It is to be noted, however, that the reports of 'pleasant' emotions were rare.) A further division was therefore undertaken in respect to those subjects who were severely disturbed (hereafter called the 'affected group') and those who were undisturbed or only mildly disturbed (hereafter called the 'unaffected group'). In the latter dichotomy no account was taken of the *kind of emotion* reported although all other information in the protocols was carefully considered to

determine the probable *intensity* of the emotional experience. Two judges who made independent classifications were found to agree in the placing of practically all subjects. In the few instances in which there was divergence of opinion, the cases were reviewed and discussed before being finally placed in one category or the other. In general, no subject who reported only 'surprise' or 'startle' was considered severely enough aroused to be included in the affected group. On the other hand all Ss who experienced some well defined emotional state in addition to surprise, unless this emotion was specifically designated as 'very mild' or 'very slight' were regarded as 'affected.'

An analysis of the classified cases in which the affected subjects are arranged according to the most *dominant* emotions which they reported is shown in Table III. The large

TABLE III

SHOWING DOMINANT EMOTIONS AROUSED (AS INDICATED BY INTROSPECTIONS)

Emotion	Stimuli				Total
	Chair	Snake	Mouse	Brains	
Anger, resentment			1	1	2
Fear, anxiety, worry, horror	7	2	2		11
Mirth	2	1		1	4
Pity			1		1
Repugnance, disgust, repulsion			1	3	4
Total affected	9	3	5	5	22
Total unaffected	9	5	3	3	20
Total cases	18	8	8	8	42

number of cases of fear, with its variations, for the chair stimulus is explained by the fact that the introspections of 10 additional subjects whose steadiness results have not yet been discussed, are here included. It should be noted that slightly more than half of the total of 42 subjects appear in the affected group, while the remainder were relatively unaffected. The fact reported by earlier investigators that the stimulus situation is no indication of the kind of emotion to be expected, is clearly demonstrated from the figures in this table.

When the contact records of the original 32 affected and unaffected Ss are separately averaged, the curves shown in Fig. 2 are obtained. The stimulus curve for the affected group shows an abrupt rise at the fifth period which is

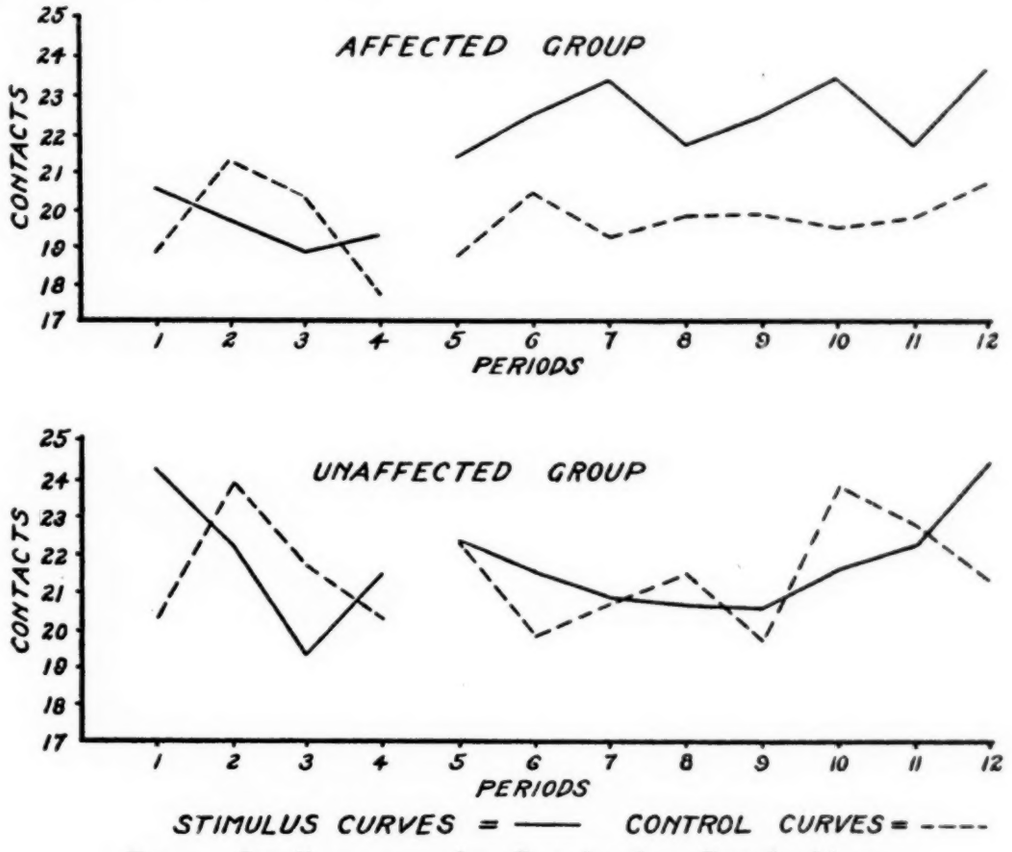


FIG. 2. Steadiness curves for affected and unaffected subjects.

maintained throughout the remaining portion of the series. The difference in the levels of the control and stimulus curves for this group is markedly greater than the difference in the levels of the control and stimulus curves for all subjects (shown in Fig. 1). The curves of the unaffected group, on the other hand, are about on the same level, and in fact may be seen to cross one another several times through their lengths. These 4 curves not only furnish objective evidence of the validity of the introspections, but also demonstrate in a striking manner, that the severely affected subjects, *regardless of the nature of their emotions*, have on the average a pronounced increase in contacts, while no such change takes place in the average steadiness of those who are relatively unaffected.

In Table IV the statistical significance of the difference between the levels of various sections of the 4 curves is computed. Thus it is seen that for the stimulus curve of the affected group, the average number of contacts per 10 sec.

TABLE IV

SHOWING RELIABILITIES OF THE DIFFERENCES BETWEEN SECTIONS OF THE 4 CURVES

	Affected Group				Unaffected Group			
	Stimulus Series		Control Series		Stimulus Series		Control Series	
	Periods		Periods		Periods		Periods	
	1–4	5–12	1–4	5–12	1–4	5–12	1–4	5–12
Ave.	19.67	22.56	19.48	19.77	21.96	21.78	21.70	21.56
N *	72	144	71	144	56	112	56	112
σ	7.21	8.51	7.83	7.65	10.37	10.14	12.00	11.01
σ_{ave}	.850	.709	.929	.638	1.32	.958	1.60	1.04
Diff.	2.89		.29		.18		−.14	
$\sigma_{diff.}$	1.11		1.13		1.84		1.91	
$\frac{D}{\sigma_{diff.}}$	2.60		0.26		0.10		−0.07	
Chances in 1000 of real diff...	995		103		40		28	

* N = the number of periods.

for periods 1–4 is 19.67, while the average number of contacts per 10 sec. for periods 5–12 is 22.56. The obtained difference of 2.89 is nearly 3 times the standard deviation of the difference, 1.11, so that the chances of this increase representing a true tendency are 995 in 1000. The differences between the corresponding sections of the other curves are, on the contrary, all unreliable.

The figures in Table V compare the reliabilities of the differences between the control and stimulus curves for periods 5–12 in both the affected and unaffected groups. Here again it may be seen that for the affected group, the obtained difference is to be considered reliable 998 times in 1000, while for the unaffected group the chances are only 64 in 1000 of the actual difference being a true one.

TABLE V

SHOWING RELIABILITIES OF THE DIFFERENCES BETWEEN CONTROL AND STIMULUS CURVES, PERIODS 5–12

	Affected Group		Unaffected Group	
	Stimulus Series	Control Series	Stimulus Series	Control Series
Ave.	22.56	19.77	21.78	21.56
N *	144	144	112	112
σ	8.51	7.65	10.14	11.01
$\sigma_{ave.}$	.709	.638	.958	1.04
Diff.	2.79		.22	
$\sigma_{diff.}$	.954		1.41	
$\frac{D}{\sigma_{diff.}}$	2.93		0.16	
Chances in 1000 of real diff.	998		64	

* N = the number of periods.

f. The effect of exercise.—In order to determine the relative influences of emotional disturbance and physical exercise upon the steadiness score, an additional group of 10 Ss whose introspections have already been classified in Table III was run. The chair drop was used as the emotional stimulus. Instead, however, of taking these subjects through the usual control series, they were required to perform a prescribed amount of physical work between the fourth and fifth periods of an additional series of tests. The work consisted in stepping upon the seat of a chair and down to the floor again 5 times in rapid succession. Since the height of the chair was 18″ from the floor and the average weight of the members of the group was 152.7 pounds, a simple computation shows that on the average they performed about 1145.25 foot-pounds of work.[11]

The effect upon the contact curve is indicated in Fig. 3, where the broken line traces the course of the steadiness when the exercise was introduced, and the solid line gives

[11] The standard physical formula,

$$\text{Work} = \text{Lift} \times \text{Load},$$

was used.

the record for the chair-drop situation.[12] The increase in contacts resulting from physical exercise appears from these curves to be of a somewhat different character than that produced by emotional excitement. Thus although unsteadi-

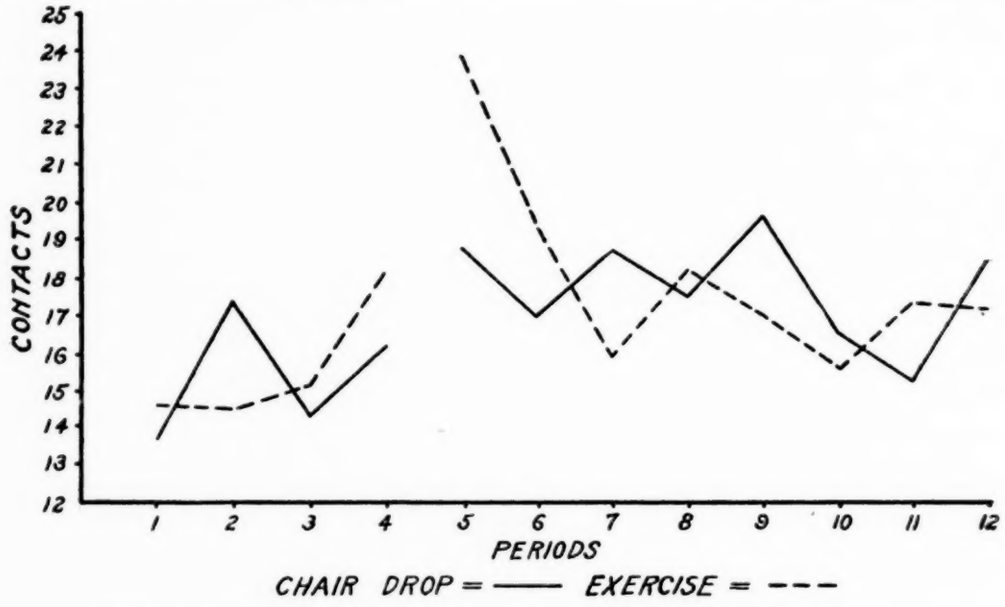


FIG. 3. Steadiness following exercise and emotional stimulation.

ness is markedly greater immediately after the work is done than it is following the chair drop, there is a slight though evident tendency for the muscular disturbance to quiet down on the average at a slower rate.

g. Interpretation.—The fact (1) that changes in steadiness score for the various emotional states are relatively uniform and (2) that the temporal pattern of these changes is different from those induced by physical exercise, may be regarded as significant. The former result moreover, finds corroboration in the diffuse and general nature of the disturbances in other bodily functions following emotional and affective stimuli. Thus Shepard (**13**) for example, has shown that feelings cannot be classified (according to Wundt's tri-dimensional theory or any other regular system) on the basis of vasomotor, heart or breathing changes; Totten (**16**) reports a common rise in metabolic rate when separate emotions are aroused,

[12] The curves presented are averaged for all 10 *Ss*. Separate curves for affected and unaffected subjects are the same in all major aspects, except that the chair drop curve for the affected *Ss* is slightly higher than the one here shown.

and Landis and Gullette (**8**) find nothing in blood pressure variations which can be regarded as characteristic of particular subjective states. The latter investigators have held in fact that "the nature and name of the emotion is determined by the external conditions rather than the reactions themselves" (**8**, p. 240). Certainly the expressive methods of measurement which have thus far been used, including those of the present study, furnish little direct evidence of a correlation between distinct kinds of organic responses and different emotional experiences. From a physiological point of view consequently it may be maintained that *there is but one emotion*, which is characterized by general disturbances in a variety of bodily activities.

A somewhat broader conception, however, which can also be supported to some extent by experimental proof would be that there are two opposite types of bodily affectations, each assuredly of a non-specific nature, which can be approximately related to different *classes* of emotions. Thus the speeding-up or hyper-activity of the functions seems to accompany what may be called the extreme or over-active states. A decrease in responsiveness on the other hand is likely to be found in conditions of subjective depression. Evidence for a distinction of this nature in abnormal cases is advanced by Henry (**5**) who holds that the apathetic, under-active and under-talkative disorders are accompanied by a retardation of the metabolic processes. The apprehensive and agitated states are correspondingly accompanied by an increased metabolic rate. Brunswick (**2**, p. 282) suggests a different though similar dichotomy on the basis of his analysis of stomach contractions following emotional stimuli. The so-called pleasant emotions, according to this author, go with an increase in the muscular tonus of the stomach and duodenum, while the unpleasant states in general are accompanied by a decrease in tonus. There are of course many exceptions to these findings in the experimental data which Brunswick reports.

What then can be concluded with reference to the present results? An examination of the emotions induced (Table III)

shows at once that no sorrowful or depressing states are among them. Indeed it is hard to see how this could have been the case with the stimulus situations employed. There was consequently little opportunity for a decrease in physiological activity even if it be supposed that such a decrease accompanies the sad or depressing emotions. Despite the arbitrary elimination of a whole class of experiences, however, one feels hardly justified in discarding entirely the introspective evidence afforded, although the disagreement among the separate statements gives weight to this proposal. The alternate position would be to consider the general increases in steadiness scores indicative of some common element or quality which is a function of all or of most all the experiences reported. It is fortunately not difficult to find such a quality. The condition of 'emotional excitement' as a matter of fact fits readily into a scheme of this sort. Thus it is quite possible to be excitedly amused as well as excitedly horrified, and to have excitement accompanying all emotional states *which are intense enough*, excepting of course those low and depressing experiences which were not elicited in this investigation. We submit therefore that the common disturbance in muscular control for the strongly affected group of subjects may be interpreted to mean that *the steadiness change is related more closely to the degree of excitement induced than to the concomitantly aroused but qualitatively different emotional states which are usually designated as 'the emotions' per se.* The natural inference which follows from this conception is that the steadiness would be found to *increase* in emotional situations of a depressing or anti-exciting kind, if indeed it could be adequately measured under such conditions. There is regrettably no means of checking this possibility from our data, although Fearing's findings (**3**) which have been previously cited lend some support to such an hypothesis.

The steadiness scores furthermore are not to be taken as valid indices of the absolute rate of muscular tremor, since the plate-and-stylus steadiness tester does not necessarily register the true tremor frequency. The increase in steadiness which was obtained in our records may in fact be due

(*a*) to an increase in the absolute tremor rate, (*b*) to an increase in the amplitude of the tremor movements, or (*c*) to both of these factors. That the changes are in no sense mere by-products of more rapid respiration in emotional situations, however, is demonstrated from the pneumographic tracings obtained in this experiment. Despite the absence of specific directions, one way or the other concerning the 'holding' of the breath, the majority of the *S*s whose breathing was recorded were found *of themselves* to inhibit breathing either partly or wholly during the test periods, and to make up for the oxygen debt thus incurred in the later rest intervals. Such behavior resulted both during the 4 emotional series and when exercise was used as a stimulus. There is to be sure some evidence that the inhibition tends to break down in the case of the more violently disturbed *S*s. For the most part, however, it appears that the steadiness changes must be traced to physiological activity less dependent upon voluntary control, of which the heart rate, incipient reflex movements of the striated musculature and vasomotor changes are possible components.

Summary

To examine the effect of emotional upset upon involuntary movements of the arm and hand a group of 42 subjects was taken through a series of steadiness tests, during which severe emotional stimuli were introduced without any previous warning. An additional series of control records without emotional stimulation were obtained with 32 of the *S*s. Steadiness was measured with an improved plate-and-stylus tester of the Whipple type. The emotional stimuli employed with different members of the group were (*a*) a chair dropped backwards through an angle of 60°, (*b*) a live snake placed on *S*'s lap, (*c*) requiring the *S* to slice a live albino mouse in two, (*d*) requiring the *S* to handle bloody calves' brains and at the same time giving him a surprise electric shock.

1. Without exception, the *S*s all reported 'surprise' or 'startle' in response to each stimulus.

2. Nearly half the *S*s experienced no other disturbance,

however, while the remainder noted one or more well defined emotional states during the period following stimulation.

3. In substantiation of the results of earlier investigators it was found that the particular stimulus used gives little indication of the kind of emotion to be expected, a considerable variety of emotions having been aroused among the different Ss by each stimulus.

4. The steadiness records treated *in toto* possess an exceptionally high degree of reliability.

5. Curves giving the degree of steadiness for successive 10 sec. periods before and after stimulation in general show an increase in contacts registered following the shock stimuli. There can hardly be said to be 'typical' curves for the different stimuli, however, nor are there any characteristic steadiness changes for the separate emotions.

6. The disturbance in muscular control on the average has been shown to continue for at least 4 minutes after stimulation, at which time the tests were terminated.

7. The influence of motivation clearly operates against a change in steadiness as signified by the fact that the subject if not too violently affected, usually tries harder to remain steady after stimulation than before.

8. There is evidence to indicate that the Ss as a whole become *more variable* in steadiness after the stimuli have been presented.

9. When the subjects are divided into 2 groups *without regard for the stimulus employed or for the emotion induced*, consisting of (*a*) those whose introspections show they are severely affected and (*b*) those unaffected or only mildly affected, a marked difference in the average steadiness curves appears. The number of contacts for the affected group is significantly greater following stimulation than before, and is also significantly greater than the control records for the same group. The average steadiness score of the unaffected group, on the other hand, remains about the same as the control record for this group.

10. After moderate physical exercise the contacts are increased more than they are by emotional stimulation.

However, the steadiness curve after exercise tends to return to its previous level more rapidly.

11. The steadiness changes in the main cannot be due to a faster breathing rate since pneumographic records show that breathing is usually inhibited during the steadiness tests.

12. A suggested interpretation, in line with previous findings upon respiratory, blood pressure and metabolic disturbances following emotional shock, is that the increase in involuntary muscular movement is probably more closely related to the *degree of excitement* induced by the emotional stimuli than to the concomitantly aroused but qualitatively different emotional states.

References

1. Blatz, W. E., The cardiac, respiratory and electrical phenomena involved in the emotion of fear. This Journal, 1925, **8**, 109–32.
2. Brunswick, O., The effects of emotional stimuli in the gastro-intestinal tone. *J. Comp. Psychol.*, 1924, **4**, 19–79 and 225–87.
3. Fearing, F. S., Factors affecting static equilibrium. An experimental study of the influence of height, weight and the position of the feet on amount of sway, together with an analysis of the variability in the records of one reagent over a long period of time. *J. Comp. Psychol.*, 1924, **4**, 91–121.
4. Garrett, H. E., Statistics in psychology and education. New York, Longmans Green, 1926, 317 pp.
5. Henry, G. W., Basal metabolism and emotional states. *J. Nerv. & Ment. Dis.*, 1929, **70**, 598–605.
6. Jastrow, J., A study of involuntary movements. *Amer. J. Psychol.*, 1892, **4**, 397–407.
7. Laird, D. A., Changes in motor control and individual variations under the influence of 'razzing.' This Journal, 1923, **6**, 236–46.
8. Landis, C., and Gullette, R., Studies in emotional reactions, III. Systolic blood pressure and inspiration-expiration ratio. *J. Comp. Psychol.*, 1925, **5**, 221–53.
9. Landis, C., Gullette, R., and Jacobsen, C., Criteria of emotionality. *Ped. Sem.*, 1925, **32**, 209–34.
10. Marston, W. M., Reaction time symptoms of deception. This Journal, 1920, **3**, 72–87.
11. Miles, W. R., Static equilibrium as a useful act of motor control. *J. Indust. Hyg.*, 1922, **3**, 316–31.
12. Moore, H. T., Laboratory tests of anger, fear and sex interest. *Amer. J. Psychol.*, 1917, **28**, 390–95.
13. Shepard, J. F., Organic changes and feeling. *Amer. J. Psychol.*, 1906, **17**, 522–84.
14. Skaggs, E. B., Changes in pulse, breathing, and steadiness under conditions of startledness and excited expectancy. *J. Comp. Psychol.*, 1926, **6**, 303–17.

15. TITCHENER, E. B., Experimental psychology. New York, Macmiilan, 1924, Vol. I, Part II, 456 pp.
16. TOTTEN, E., Oxygen consumption during emotional stimulation. *Comp. Psychol. Monog.*, 1925, 3, No. 13, 79 pp.
17. WHIPPLE, G. M., Manual of mental and physical tests. Baltimore, Warwick and York, 1924, I, 367 pp.

(Manuscript received February 2, 1931)

EFFECT OF PITCH OF TONE-STIMULI UPON BODY RESISTANCE AND CARDIO-VASCULAR PHENOMENA

BY L. E. MISBACH[1]

I. Statement of Problem

Many experimenters have noted that auditory stimuli may evoke circulatory changes and changes in body potential or body resistance to electrical current. These studies have, however, involved meaningful or startling sound-stimuli rather than tones. The major purpose of the experiment reported here was to study galvanic and cardio-vascular responses as determined by variations in pitch of tone-stimuli when loudness is constant. The responses sought were those which might be presumed to be correlated with discomfort or displeasure occasioned by the stimuli. The stimuli were, therefore, made too loud to be soothing or pleasing but not loud enough to occasion pain, and were not of such duration as obviously to cause fatigue. Startle or shock effect was, as far as possible, eliminated by presenting each tone at very low intensity, bringing it slowly up to the loudness desired. As two loudness levels of stimuli were used, the experiment also includes a limited study of responses as determined by variation in loudness when frequency is constant.

II. History

Fere and Londe (1) were among the earliest observers of responses to tone-stimuli. They studied changes in dynamometric pressure occurring with changes in pitch and intensity of tone stimuli produced by tuning forks. They used frequencies ranging from 50 to 1000 v.d. Their conclusion is that muscular tension increases with pitch and intensity of auditory stimuli.

DeRochas (2) presented isolated tones to a hypnotized subject. He found that this subject responded with a wide-spread trembling to notes extremely high or extremely low in pitch.

More recently, Mrs. Haller (3) has observed reactions of infants to tone-stimuli.

[1] A study from the Psychological Laboratory of Northwestern University carried out under the direction of Dr. John J. B. Morgan.

Some of Mrs. Haller's conclusions are: Increase in intensity of pure tones is directly related to disturbed-responses of infants. Higher pitches are more disturbing than low. With continued stimulation, increase in frequency is more effective than increase in intensity in evoking signs of displeasure.

Laird and Coye (4) report an experiment in which pure tones were compared for conscious annoyance. The stimuli were produced by the Western Electric 2A audiometer. In a part of the experiment; paired comparisons of annoyance quality were made with tone-stimuli of eight octaves, from 64 to 8196 v.d., at sensation level 50. Results of this work indicate that at equal sensation level of 50, frequencies below 256 v.d. and above 512 v.d. are more annoying than tones ranging from 256 to 512 v.d.

Wechsler (5) found that strong auditory stimuli evoke the galvanic response. He held that sensation, to be effective in evoking the galvanic response, must be of such a nature as to produce an emotional state—surprise, pain, or displeasure. He concluded also that the magnitude of the galvanic response (exsomatic current being constant) is roughly proportional to the intensity of the affective state evoked.

Darrow (6) has also used auditory stimuli to evoke galvanic responses. He reports that changes in apparent resistance and body potential are consistently greater to sensory stimuli than in response to ideational stimuli. Auditory stimuli used by Darrow were for the most part such as to startle or shock the subject.

Darrow, in the same experiment, found an average rise in both blood pressure and heart rate following sensory stimuli, but these changes were small as compared with those occurring in response to ideational stimuli. In another article (7) he notes that Zoneff and Meumann and Brahn found a faster pulse rate in unpleasantness or pain, a slower rate in pleasantness.

III. Apparatus

Psychogalvanic responses were recorded from deflections of the Hathaway galvanometer. The electrode current with the Hathaway is a 60 cycle A.C. The current passing through the electrodes is kept constant during operation by adjustment of a potentiometer, which regulates the voltage imposed on the subject to balance slow changes in the resistance of the body of the subject. Zinc electrodes three centimeters in diameter were used, being applied with a paste of water and seven parts kaolin to one part zinc sulphate.

The deflections of the detecting milliammeter were recorded photographically. For this purpose a photographic device perfected by Dr. A. R. Gilliland was used. A roll of six-centimeter bromide paper in a photographic box is driven by a motor and a gear mechanism so as to be exposed to beams of light passing through a slot in the box. A second milliammeter placed in front of this slot is connected in series with the detecting milliammeter of the Hathaway instrument. A small mirror is supported by a light bronze wire which is attached to the needle of this second milliammeter in such a way as to be directly over and parallel with the axis of the hand. Compensation for this change in balance of the needle has been made by adjustment of the coil of the milliammeter. A beam of light from a 25 c.p. bulb is focussed so that it is reflected from this milliammeter mirror through the slot upon the moving roll of bromide paper in the photographic box.

Blood pressure and pulse rate were also recorded photographically on the bromide paper. The tube of a blood pressure cuff forks, one part being attached to a mercury manometer, the other to a metal tambour. The tambour motivates a lever, to the axis of which is attached a small mirror. A beam of light is reflected from this mirror through the slot so as to fall upon the photographic paper.

A base line was provided by a third light focussed from a small stationary mirror. A telegraph key was used to operate an electromagnet, the lever of which intercepted the base-line light. By this means the end of each stimulus was marked. As it produced an audible click, it was not used to indicate the beginning of a stimulus for fear that an anticipation response on the part of the subject might result.

The stimuli were produced by the Western Electric 2B audiometer. In order to secure the desired intensities, the tones were given radio amplification (8). The frequencies given by the audiometer are 64, 128, 256, 512, 1024, 2048, 4096, and 8192 v.d. Each tone is produced at will by an individual key. An attenuator dial is calibrated in sensation units, so that by turning the dial the desired sensation level of any one of the eight tones above can be produced directly from the audiometer. The subject received the stimuli through ear-phones. This was found a practical necessity in order to be sure of a standard of loudness throughout the experiment.

A Western Electric oscillator and a Western Electric 34A amplifier were used in standardization of the stimuli.

The tests were conducted in a quiet room. A cot was provided upon which the subject half-reclined in a comfortable position. A screen was placed so as to conceal apparatus and operators from the subject.

IV. Standardization of Stimuli

It was desired in this experiment to use tones ranging in frequency in steps of octaves from 64 to 4096 v.d. in two levels of loudness. By loudness is meant a purely psychological aspect of sound. Fletcher (9) defines loudness as the magnitude of sensation produced by a sound. Whereas loudness is a psychological factor, intensity of sound is the physical energy of the sound waves in air or other media. Loudness of sound does not correlate closely with either intensity or sensation level. Tones of equal intensity or of equal sensation level but different in pitch are not equal in loudness. For extreme differences in pitch, this inequality is great. Equality of loudness is, of course, based on a hypothetical average ear. Fletcher summarizes an experiment by Kingsbury in the Bell Telephone Laboratory, the purpose of which was to equalize the loudness of tones. Judgments of twenty-two subjects were averaged to establish equality of loudness of tones of a wide range of pitch.

In order to obtain the loudness desired for this experiment, it was necessary to amplify the audiometer tones. The relative effect of the amplifier on the different frequencies and the range of efficiency of the ear-phones used were not known with sufficient exactness to make it possible to obtain from the audiometer dial the sensation levels or intensity of the tones produced. It was therefore necessary to obtain subjective judgments to establish equality of loudness of the tones desired.

In obtaining a measure of equal loudness of stimuli, Kingsbury's methods were, as far as possible, copied. A Western Electric oscillator and 34A amplifier were used to produce a standard tone of 700 v.d. at the loudness desired. This standard tone was produced through an ear-phone and was kept constant throughout the work of obtaining judgments for each level. The tones to be standardized for stimuli were produced through the ear-phones which were used throughout the whole experiment. The dial of the radio amplifier was calibrated to permit more delicate attenuation of the tones to be standardized than could be obtained by the audiometer dial only. In making judgments the subject held in one hand the earphone in which the standard tone was produced and in the other hand held one of the ear-phones in which the tone to be standardized was produced. He then placed the two ear-phones to his ear in rapid alternation, directing the operator to change the audiometer tone up or

down as the subject judged necessary to make it equally loud with the standard tone. This continued until the judgment 'equal' was given, the dial readings being then recorded. Three judgments for each tone were made by each subject. The tones were presented in chance order, two judgments of the same tone seldom being made consecutively. Judgments of fifteen subjects were averaged and the dialings thus obtained for the stimulus-tones were used throughout the later experimental work.

It was found that the tone of 4096 v.d. could not be strengthened by amplification, possibly being beyond the range of maximum efficiency of the ear-phones used. It was also found that the maximum loudness obtainable for the tone of 64 v.d. was not as loud as the upper loudness level desired. Judgments were made, therefore, for only five tones (128, 256, 512, 1024, and 2048 v.d.) in the upper loudness level, and for six tones (64, 128, 256, 512, 1024, and 2048 v.d.) in the lower loudness level. The tone 4096 v.d. was used as a stimulus, but was much less loud than the lower level of loudness of the other stimuli.

In order to have some objective measure of loudness, Fletcher (10) suggests that loudness be defined as the sensation level of a reference tone of 1000 v.d. The loudness of any tone is expressed as the sensation level of a tone of 1000 v.d. which is judged to be equal in loudness with the tone to be measured. On the basis of Kingsbury's work, Fletcher gives graphs showing the relation of loudness, as defined above, of different frequencies to sensation level and intensity level.

An objective estimate of the loudness of the tones was obtained by comparing for loudness the standard tone at the two loudnesses used with a tone of 1024 v.d. produced without amplification by the audiometer, which gives readings in sensation level. The tone 1024 v.d. reaches a greater loudness than any other on the audiometer, and it was found that this loudness is slightly greater than the upper level of loudness established for the stimulus-tones.

The loudness, intensity, and sensation levels of the stimuli used in this experiment were, by this means, estimated to be as follows:

TABLE I

LOUDNESS, INTENSITY, AND SENSATION LEVELS OF STIMULI

Upper Loudness

Frequency	128 v.d.	256 v.d.	512 v.d.	1024 v.d.	2048 v.d.
Loudness level	98	98	98	98	98
Intensity or phonic level in decibels	+ 12	+ 9	+ 7	+ 5	0
Sensation level	78	87	91	98	94

Lower Loudness

Frequency	64 v.d.	128 v.d.	256 v.d.	512 v.d.	1024 v.d.	2048 v.d.	4096 v.d.
Loudness level	87	87	87	87	87	87	70
Intensity or phonic level in decibels	+ 2	0	− 2	− 3	− 5	− 6	− 24
Sensation level	52	62	75	85	87	87	70

Loudness level is expressed as the sensation level of an equally loud tone of 1024 v.d. Sensation level is the number of sensation units of a tone above its threshold of audibility.

It was desired that the stimuli should be as near pure tones as possible. Amplification may, however, have introduced overtones, especially in the lower frequencies. These overtones, if present, were not detected by the subjects who standardized the stimuli nor by later subjects.

V. Procedure and Results

A total of 29 subjects were given the stimuli described above. They were graduate and undergraduate students in Northwestern University, ranging from twenty to thirty years of age. Of the total number, 13 were men and 16 women. No subject with known hearing abnormality was used.

The series of twelve stimuli was presented only once to each subject. The order of presentation of stimuli for each individual test was determined by shuffling a pack of cards, so that the stimuli were never given in similar order to two subjects.

As it was desired to obtain responses to the character of the tonal sensation, suddenness of presentation of stimuli was eliminated as far as possible. It seemed probable that, if the stimuli had shock value, responses might be relatively uninfluenced by the quality of the stimulus following its initial shock. Stimuli were, therefore, presented in the following manner. The amplifier dial was turned back so that no tone would come through, and the audiometer dial was set at the position required for the stimulus to be given. The operator threw in the key to the tone desired, at the same time with the other hand turning the amplifier dial slowly upward. Throwing in the key removed a shutter from the base-line light, thereby giving a record on the film of the time of presentation of the stimulus. After the audiometer key was thrown in, five seconds were allowed to bring the tone to its maximum. It then remained at maximum for ten seconds, when the operator turned the amplifier back, throwing the key out when the dial had been returned to the point of inaudibility. About three seconds were allowed to dial the sound stimulus out of audibility. Starting and stopping suddenness was thus reduced to the minimum. No clicks or other sounds were allowed to occur to warn the subject that a stimulus was about to be given. The ten-

second duration at maximum of the stimulus was suggested by Mrs. Haller (11) as allowing time for any physiological reaction to be aroused, without being so long as to bring in the element of fatigue.

A fairly routine order was observed in making the tests. The subject was asked to lie on the cot in a comfortable position. He was told that auditory stimuli would be presented, and assured that, while some of them would be loud, none would be painful. The electrodes were smeared with the kaolin and zinc sulphate paste and were bound on the palms of the hands, and the galvanometer was placed in circuit. The blood-pressure cuff was then wrapped about the ankle and inflated to 80 mm. Hg. The lights were then focussed on the slit in the photographing box. The blood-pressure cuff was deflated for a minute or so and the head-phones were adjusted on the ears of the subject. The blood-pressure cuff was again inflated and the motor of the photographing box started. The galvanometer needle was steadied near the zero point by turning the potentiometer as needed before the first stimulus was given.

The stimuli were presented in chance order with a minute interval between successive stimuli. If there was no observable galvanic response to a stimulus, a full minute interval before presentation of the next stimulus was not always allowed. If, on the other hand, a large deflection occurred, time was allowed to permit the return of the needle to near the zero point. If it did not return, the potentiometer was used to return it near the zero point again.

The blood pressure was kept at about 80 mm. Hg. as this was judged to be well above diastolic pressure in the tibial artery and on the ankle did not give much discomfort during the time required for the test. In a few cases the cuff was deflated for a brief period after half the stimuli had been given. Usually it remained inflated during the presentation of all twelve stimuli.

Blood-pressure changes were translated from the photograph into mm. Hg., the relation between movements of the light on the film and the movement of the mercury having

been experimentally determined. The changes were read as deflections from the prevailing base line of the pulse beat. Only changes occurring between the beginning and end of the stimulus, as shown by the break in the base-line light record, were counted, as later changes may or may not be due to the stimulus.

Galvanometer deflections were recorded in millimeters of the maximum height of the deflection of the galvanometer line on the film. This height was measured from the prevailing base line of the galvanometer line on the film. Only deflections beginning within the duration of the stimulus were counted. Averages of actual magnitudes of galvanic responses are, however, of little or no value because different subjects show widely different orders of galvanic response. Moreover, although the electrode current was kept constant by adjustment of the potentiometer, the actual magnitude of current was unknown. To obtain a measure of average magnitudes, galvanic deflections were scored on a percentage basis. The largest response of an individual record was taken as 100 percent for that person. All responses for each individual are thus recorded on a scale, the largest number of which is 100 and the smallest possible, zero. Thus responses of different individuals affect the final average regardless of the different orders of actual magnitude of response.

The galvanic responses occurring were, almost without variation, an apparent lowering of body resistance. Heightened resistance occurred occasionally as a small initial change of not over two seconds duration preceding a lowered resistance. These infrequent small increases in resistance were ignored in recording the data.

Pulse-rate changes were recorded as the differences, plus or minus, between the ten seconds of duration of the stimulus at its maximum strength and the ten-second interval immediately preceding presentation of the stimulus. In the record, plus indicates an increase in rate, minus a decrease.

Eighteen subjects showed blood-pressure responses; fourteen showed pulse-rate changes; and twenty-three showed

galvanic responses. A summary of all data obtained from the test is given in Table II.

TABLE II

SUMMARY OF ALL RESPONSES OF 29 SUBJECTS

Lower Loudness Level

I Stimulus	II Total No. Blood-Press. Responses	III Total Magnitude of Blood-Press. Resp. (mm. Hg.)	IV Total No. Pulse-Rate Changes	V Total No. Galvanic Responses	VI Average Percent Magnitude of Galv. Resp.
64 v.d.....	5	.15 −.09	2	18	49
128 v.d.....	3	.22	4	19	48
256 v.d.....	4	.6 −.12	0	14	30.8
512 v.d.....	4	.13 −.13	1	17	38.5
1024 v.d.....	6	.32 −.06	2	16	37
2048 v.d.....	4	.43	3	19	38.5
4096 v.d.....	4	.16 −.13	1	19	42.2

Upper Loudness Level

I Stimulus	II Total No. Blood-Press. Responses	III Total Magnitude of Blood-Press. Resp. (mm. Hg.)	IV Total No. Pulse-Rate Changes	V Total No. Galvanic Responses	VI Average Percent Magnitude of Galv. Resp.
128 v.d.....	8	.59 −.10	6	20	60.5
256 v.d.....	5	.34 −.03	2	20	49
512 v.d.....	5	.15 −.10	2	16	32.1
1024 v.d.....	5	.29 −.10	4	20	43.4
2048 v.d.....	5	.18 −.26	3	22	63.5

Stimuli are listed in vertical column I. The total number of blood-pressure responses to each stimulus is given in column II; the total magnitude of blood-pressure responses to each stimulus in mm. Hg. is given in column III. Column IV shows the total number of pulse-rate changes in response to each stimulus. Column V gives the total number of galvanic responses (lowering of body resistance) to each stimulus; and vertical column VI, the average percentage of magnitude of galvanic responses to each stimulus, 100 percent being the greatest deflection in any individual record and 0 percent indicating no response. (Detailed data are given in thesis, "Effect of Intense Auditory Stimuli upon Body Resistance and Cardio-Vascular Activity," filed in Northwestern University Library, August 1930.)

Blood-pressure Responses

The records show a total of 59 blood-pressure changes. This is about 14 percent of the total number of stimuli presentations. Of this total of 59, 46 represent an increase in blood pressure, 13 a decrease. Of 18 blood-pressure changes occurring with pulse-rate changes, 14 are increases in pressure accompanied by increase in pulse rate. Fifty-two of the blood-pressure responses occur with galvanic responses. As there are only 220 galvanic responses to 348 stimuli, this coincidence is significant.

No pattern of frequency effects could be considered legitimate with so few responses, as chance factors must be supposed to enter largely into their distribution. It may be noted, however, that the average number of blood-pressure responses occurring with stimuli of the upper loudness level is 5.8 and the average magnitude, .07 mm., while the average number with stimuli of lower loudness is 4.5 and the average magnitude, .062 mm. The largest number occurs with frequency 128 at the upper loudness level. Blood-pressure responses are found to be broadly in harmony with the pattern presented by galvanic responses.

Pulse-rate Changes

A total of thirty changes in pulse rate are recorded. The pulse rate of 7 records, however, could not be counted because of fault in the time line. Twenty-one responses are increases in pulse rate, 9 are decreases. Eighteen of the total number of 30 pulse-rate changes coincide with blood-pressure changes. Twenty pulse-rate changes occur with galvanic responses. The largest number of pulse changes is recorded with the upper loudness level.

Galvanic Responses

To a total number of 348 stimuli presentations there are 220 galvanic responses. Six subjects showed no galvanic response to any stimulus. The total numbers and total and average magnitudes of galvanic responses show a consistent pattern of frequency and loudness effect. There are a larger

number and larger average magnitude of responses in the upper than in the lower loudness level. There are also consistently larger numbers and magnitudes of responses to tones of frequency 64, 128, 2048, and 4096 than to 256, 512, and 1024.

Adaptation

There is, however, an apparent effect of adaptation which affects the validity of totals and averages in percentage, as well as in actual magnitude of galvanic deflection. It was obvious upon even a casual study of the records that there is a pronounced tendency for responses to be at their maximum for the first few stimuli presented, regardless of the nature of the stimuli, and for the responses to fall away toward zero as the test continues. Such effects are, it seems, a common phenomenon in work with the galvanometer, both polarization and subjective factors being usually considered operative. As the exsomatic electrode current used in this experiment was a 60 cycle alternating current and as the actual current between the electrodes was kept roughly constant, polarization would seem to be of less significance than subjective factors in determining adaptation effects.

In order to estimate the effect of adaptation on the galvanic response, the percent of magnitude of all galvanic responses was recorded according to the order in which they occurred, whether first, second, third, etc., in each individual test, without regard to the nature of the stimuli. All the responses occurring first in the individual tests were then averaged, and those occurring second, third, etc., were likewise averaged. These averages are given in Table III.

TABLE III

ADAPTATION EFFECTS

Averages of percentage of magnitude of galvanic responses according to the order in which they occurred in individual tests, disregarding nature of stimuli. (Complete data may be found in thesis.)

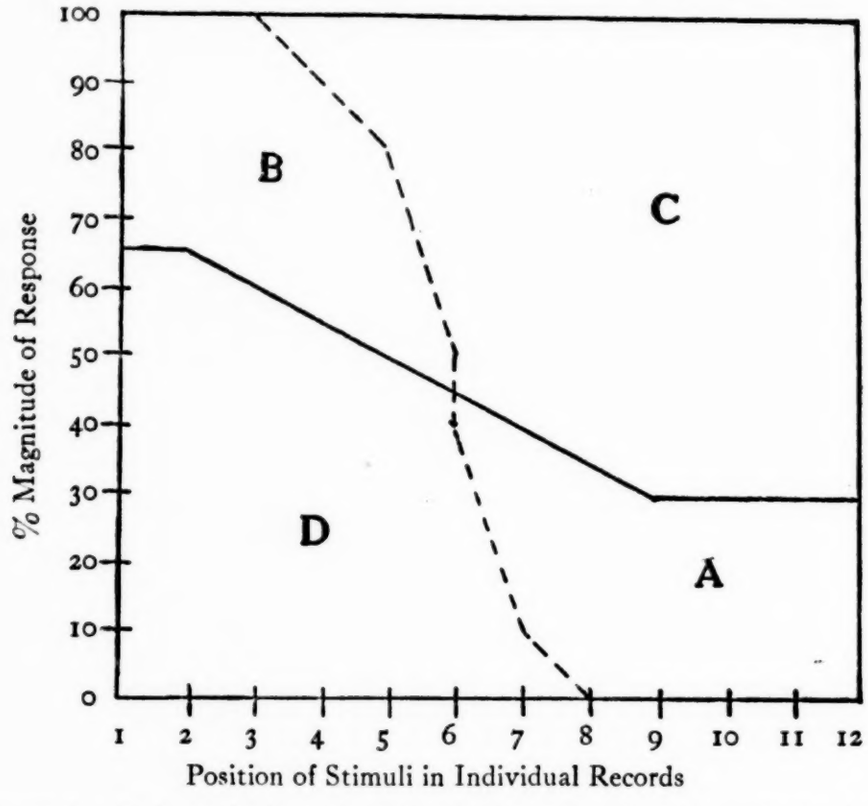
Position	1	2	3	4	5	6	7	8	9	10	11	12
Average magnitude....	67.5	65.8	54.6	56.8	53.1	45.1	39.4	33.3	19.2	28.1	37.6	35.9

It seems probable that the effect of adaptation should be represented by a smoother curve than given by the actual averages. There is no obvious reason, for example, for a sudden drop between positions 8 and 9. The averages are, therefore, smoothed out to read: 65, 65, 60, 55, 50, 45, 40, 35, 30, 30, 30, 30. These averages are represented by the unbroken line in Graph I. It is likely that the average effect of adaptation is more accurately represented by a curve as indicated roughly in Graph I, than by a straight line. As adaptation may be correlated with expectation of pain or similar subjective factors, there is no reason to suppose that it should show a constant ratio of diminution of response.

It is obvious, also, that the effect of the variations in quality of stimuli would be to pull the graph line away from the diagonal; that is, if stimuli of equal response value were presented instead of stimuli of different response values, the adaptation line would run more nearly from one hundred to zero. It may, therefore, be assumed that responses for position 1 would, if adaptation were the only effect, run from 65 to 100. On the other end of the scale, it may be assumed, likewise, that the effect of adaptation would throw the response between thirty and zero. There would be a median point at which responses to adaptation effects would fall on either side of the unbroken line. Because of the curvilinear character of the unbroken line, this median point is placed at 6. In order to graph the range of probable adaptation effects, the broken line is drawn in Graph I across the unbroken line. The end points of this line are, as stated above, one hundred and zero, respectively. The broken line must cross the unbroken line at 6 as explained above. Other points on the broken line are determined by an attempt to make spaces *A* and *B* in Graph I of approximately equal area. The resulting graph, as stated above, is to be regarded as an estimate of the range of typical adaptation effects if stimuli were all of equal response value.

Graph I is used here as a test of the validity of the averages obtained in Table II. It is possible, for example, that a determining number of responses to frequency 128 or 2048

might be large because of occurring early in individual tests rather than because of the actual character of the stimulus. It must be remembered, however, that frequency or loudness



GRAPH I. Estimated adaptation effect in determination of magnitude of galvanic responses. The numerals on the abscissa indicate the order in which stimuli were presented, disregarding quality of stimuli. The percentage of magnitude of galvanic response is indicated on the ordinate. The unbroken line indicates the estimated minimum effect of adaptation in determining magnitude of responses. The broken line indicates the maximum effect of adaptation. Responses of such position in any individual test and of such magnitude as to fall in areas *A* or *B* are presumed to be possibly determined in magnitude by adaptation effects. Responses falling in areas *C* or *D* are presumed to be determined in magnitude by factors other than adaptation.

effect and adaptation effect may work together as often as in opposition. Responses in areas *C* and *D* are judged to be independent of probable adaptation effect, while in the case of those falling in areas *A* and *B*, the effect of adaptation cannot be differentiated from stimulus-quality effects. From the total of 220 galvanic responses 98 were found to fall in areas *C* and *D*. Table IV gives the average percentage of magnitude of these responses.

TABLE IV

AVERAGE MAGNITUDES OF GALVANIC RESPONSES OCCURRING INDEPENDENT OF ADAPTATION

Lower Loudness Level

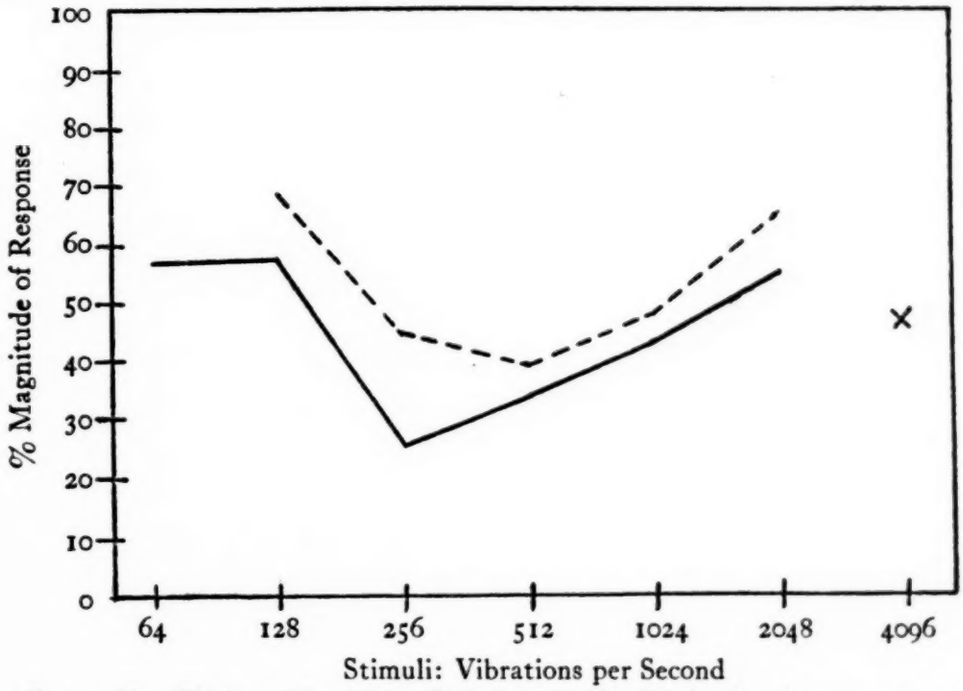
Stimuli	64 v.d.	128 v.d.	256 v.d.	512 v.d.	1024 v.d.	2048 v.d.	4096 v.d.
Average magnitude of galvanic response.....	56.4	57	25	33	42	54	48

Upper Loudness Level

Stimuli	128 v.d.	256 v.d.	512 v.d.	1024 v.d.	2048 v.d.
Average magnitude of galvanic response.....	68	44	37.9	46.6	65

Under each stimulus is shown the average percent of magnitude of the 98 galvanic responses which are independent of adaptation effects. Only responses falling in areas *C* and *D* of Graph I are included in these averages. (For detailed data see thesis.)

These averages are consistent with the pattern shown by averages of Table II. They are probably the most reliable measure of effects of stimulus variation. Graph II, below,



GRAPH II. Pitch and loudness effects on magnitude of galvanic responses. The different tones used as stimuli are indicated on the abscissa in number of vibrations per second. Percentage of magnitude of galvanic responses is indicated on the ordinate. The unbroken line shows the average responses to tones in the lower loudness level (estimated as equal in loudness to a tone of 1024 v.d. at 87 sensation units above its threshold of audibility). The average for tone 4096 is indicated by *x*, as it was nearly 20 sensation units lower in loudness than other tones of the lower level. The broken line shows average responses to tones in the upper loudness level (estimated equal to 1024 at 98 sensation level).

is based on the averages obtained in Table IV and shows final estimated frequency and loudness effects on galvanic response.

VI. Discussion

It seems probable that in previous work with galvanic and vascular responses to stimuli, sensory discomfort has been only one of many factors in determination of responses. Darrow (**12**) offers the theory that galvanic response to sensory stimuli is concomitant with lack of ready neural organization to make possible an easy or habitual response. This suggests that the magnitude of galvanic response might depend on the degree of neural and muscular activity in an attempt to make escape or other adequate response, or upon the extent and character of the ultimate or habitual response. Factors other than displeasure which have commonly affected response to sensory stimuli are startle or shock effect, meaning, and emotional value. Suddenness may well produce the effect suggested by Darrow. A sound with meaning might be expected to produce effects according to its association in experience. Some sounds, such as a gunshot, a fierce growl, or any sudden, intense sound, might well cause a reflex fear reaction. The effect of shock was reduced in this experiment by introducing the stimuli at low intensity. Expectation of pain was also avoided as much as possible by assuring the subject that painful stimuli would not be given. As very few meaningful sounds in common experience are even relatively pure tones, and as the subject knew in advance that the sounds presented would be merely loud tones, it is unlikely that meaning could have been important in determination of responses. It seems reasonable to consider responses in general determined by degree of displeasure given by the stimuli.

No valid frequency-determined pattern for blood-pressure and pulse-rate responses is presented because so few responses occurred. The average magnitude of the blood-pressure responses recorded is .066 mm. Hg. This is of a much smaller order or magnitude than are blood-pressure responses usually obtained when startling stimuli are presented.

Pulse-rate changes are, likewise, much fewer and smaller in magnitude than are the changes commonly produced by sensory stimuli.

Sixty-four percent of the stimuli presented elicited galvanic responses. In the upper loudness level 67 percent of the stimuli presented produced galvanic responses; in the lower loudness level, 58 percent were effective. A consistent difference of ten percent average magnitude of galvanic responses for the two loudness levels is shown in Graph II. This suggests that at lower loudnesses a point would be reached at which few, if any, galvanic responses would be elicited by sensory displeasure alone.

The evidences noted above indicate that, while unpleasant sensory stimulation is a factor in noise effect, other factors such as meaning, conditioning of the subject, or fatigue are generally of greater importance. It is probable that brief stimulations which are most unpleasant would produce greater fatigue results than less unpleasant stimuli. The proof of this probability is, however, beyond the scope of the tests summarized here.

It has usually been found that behavior and verbal report indicate that the disturbing power of tone increases rapidly above 1024 v.d. Laird (**13**) found, however, that the least disturbing frequencies are 256 and 512 v.d. He suggests that it may be significant that this corresponds to the normal range of the human voice. Laird notes also that the greater relative effect of high and extremely low frequencies becomes important in view of the fact that few sound-insulating materials are made to eliminate these frequencies.

It is possible that the large average of responses to tones of 64 and 128 v.d. in this test are related to their nearness to the threshold of feeling; that is, the intensity at which pain or 'fluttering' is produced by sound. According to Fletcher's estimates, 128 v.d. at loudness level 98 is about 18 decibels below the feeling threshold, while 64 v.d. at loudness level 87.5 is about 12 decibels below the threshold of feeling. As 2048 v.d. at loudness level 98 is at least 30 decibels below the feeling threshold, it is unlikely that this should be a factor in results obtained for the upper frequencies.

Whether or not the pattern of responses obtained for these loudness levels will be found the same in lower levels of loudness can be ascertained only by further experimentation. As the pattern for the higher frequencies agrees with results obtained by other methods, it appears likely that the upper half of the pattern will remain somewhat the same in all levels.

By comparing Graph II with the relationships between estimated loudness, intensity, and sensation level in Table I it is possible to estimate the effect of equal intensity and equal sensation level on response-pattern within the loudness limits of the stimuli used in this test. At equal sensation levels in the upper range of loudness, frequencies below 256 v.d. should show a greater disturbing power than median frequencies; and this difference would be greater than in equal loudness levels, since the lower pitches would gain faster than the median pitches in loudness. Frequencies about 1024 v.d. should show a greater disturbing power than median frequencies, and the pattern of differences would remain the same as in equal loudness levels up to 4000 v.d., since in this range equality of sensation level corresponds with equality of loudness. As equal loudness represents a fall in intensity from the low to the high frequencies, at equal intensity level in a range from − 10 to 10 decibels of intensity, tones below 256 v.d. should show little or no increase in disturbing power over tones of the median range. The pattern would tend to smooth out into equality as far as 1024 v.d., above which disturbing power would increase much more rapidly than it increases with frequency when loudness is constant (**14**).

VII. Conclusions

1. Brief tone-stimuli of relative purity and not of such loudness as to give actual pain do not, when presented to normal adult subjects, evoke a significant number or magnitude of blood-pressure or pulse-rate changes, if shock and expectation of pain are eliminated.

2. Equally loud tone-stimuli, which are loud enough to be unpleasant but not so loud as to give pain, and which do not surprise, frighten, or fatigue the hearer, will usually elicit lowering of body resistance to electrical current; and

the probability of occurrence and the magnitude of these resistance changes is directly related to the pitch of the stimuli. Galvanic responses to tones of frequency higher than about 512 v.d. occur more often and increase in average magnitude as frequency becomes higher. Galvanic responses to tones of frequency lower than about 256 v.d. occur more often and increase in average magnitude as frequency becomes lower. Tones ranging in frequency from about 256 to 512 v.d. elicit fewer and smaller galvanic responses than tones of frequency either above or below this range.

3. Judgment of equal loudness of tones is not a measure of their disturbing power as indicated by galvanic response.

4. Within a loudness range equal to the loudness of a tone of 1024 v.d. at 87 to 98 sensation units above its threshold of audibility, a change of one octave in pitch above 512 v.d. or below 256 v.d. produces an average change in the magnitude of galvanic response greater than is produced by a change of 10 decibels in intensity in any frequency.

References

1. Fere, Chas., *Sensation et mouvement* (Paris: F. Alcan Co., 1887), pp. 36–41.
2. DeRochas, A., Les sentiments, la musique, et la geste, *Nouvelle Revue*, CXIV (1898), 381–9; 586–602.
3. Haller, Mary W., *The reactions of infants to changes in the intensity and pitch of pure tones*, Master's Thesis, Northwestern University, 1929.
4. Laird, D. A., and Coye, Kenneth, Psychological measurements of annoyance as related to pitch and loudness, *Journal of the Acoustical Society of America*, October, 1929.
5. Wechsler, D., The measurement of emotional reactions: researches on the psychogalvanic reflex, *Archives of Psychology*, XII (1925), 188 ff.
6. Darrow, C. W., Electrical and circulatory responses to brief sensory and ideational stimuli, *Jr. of Exper. Psychol.*, XII, No. 4 (August, 1929).
7. Darrow, C. W., Differences in the physiological reactions to sensory and ideational stimuli, *Psychol. Bull.*, 1929, p. 26.
8. The amplifier unit of the Majestic Model 100 radio was used. This apparatus was furnished by courtesy of the Grigsby-Grunow Company, Chicago, Ill.
9. Fletcher, Harvey, *Speech and hearing* (New York: D. Van Nostrand Co., 1929), p. 226.
10. Fletcher, Harvey, *loc. cit.*
11. Haller, Mary W., *op. cit.*
12. Darrow, C. W., Electrical and circulatory responses to brief sensory and ideational stimuli, *Jr. of Exper. Psychol.*, XII, No. 4 (August, 1929).
13. Laird, D. A., and Coye, Kenneth, *op. cit.*
14. Fletcher, Harvey, *op. cit.*, p. 230.

(Manuscript received March 28, 1931)

THE FORM OF THE CURVE OF MEMORIZING

BY GOODRICH C. WHITE[1]

Emory University

In a study of the effect of continuous memorizing upon the rate of memorizing itself and upon retention (**6**), opportunity was afforded to study the form of the curve of memorizing under three conditions: 1, in the course of practice in memorizing; 2, in the course of a rather prolonged period of continuous memorizing; and 3, in the relearning of material memorized twenty-four hours previously, during the period of continuous memorizing.

In the experiment twenty-five subjects first went through a practice series in which they memorized, in varying orders, twenty-five lists of twelve nonsense syllables each. The lists were learned on successive days, two or three lists per day, with rest pauses between each list. Twenty-four hours after the end of the practice series the subjects memorized twenty-five lists at a single sitting, the work being without rest between lists, and requiring from two to five hours. Twenty-four hours later the subjects relearned five of the lists memorized during this period of continuous work: the first, the seventh, the thirteenth, the nineteenth, and the twenty-fifth.

The anticipation method was used, with a two-second exposure of each syllable. Throughout the course of the experiment record was kept of each correct anticipation in each successive presentation.

To make possible the comparison of curves based upon varying numbers of repetitions or presentations of lists the Vincent method of plotting the curves was adopted. This method, first used by Vincent (**5**) for the comparison of curves of maze learning by animals, has been subsequently used by Kjerstad (**1**) and by Robinson and Heron (**4**) and Robinson and Darrow (**3**) in the comparison of curves for

[1] From the Laboratory of the University of Chicago.

memorizing. Briefly the method consists in finding what percentages of material are learned in certain selected percentages of the total time required for the complete learning of that material. In this study the unit of 'time' is a single presentation of the list, and the percentages of material memorized in successive quarters of the total time were computed.

The question specifically becomes then: Is there, under the varying conditions of the experiment, any significant variation in the relative amounts learned in successive quarters of the total time required for learning the single lists? The following were chosen for comparison: the lists in positions 1, 7, 13, 19, and 25 of the practice series; those in positions 1, 7, 13, 19, and 25 of the continuous work series; and the five lists relearned by each subject, these being the same as those for which curves are plotted for the continuous work series.

The number of correct anticipations at the end of each quarter of the total learning time was computed for each list for each subject in each of these positions. For example, if a subject is able to anticipate all the syllables of a list correctly on the seventeenth presentation, he has required sixteen presentations for its mastery. The quartile points would then come at the end of the fourth, the eighth, and the twelfth presentations, and the number of correct anticipations on the fifth, the ninth, and the thirteenth presentations would represent the amount learned in the successive quarters. In cases where the quartile point falls within a presentation, the proportionate part of the increment for that presentation is added to the score for the preceding presentation. For example, if the total number of presentations required is fifteen, the quartile points then fall at 3.75, 7.5, and 11.25. Now if the number of correct anticipations after three presentations (*i.e.*, in the fourth presentation) is four and the number of correct anticipations after four presentations (*i.e.*, in the fifth presentation) is five, the correct score for the first quartile point is taken as 4 plus (.75 × 1), or 4.75. Likewise if after seven presentations the score is eight correct antici-

pations and after eight presentations the score is ten correct anticipations, then the fifty percentile point score is taken as 8 plus (.50 × 2), or 9.

The scores thus determined for successive quartile points for each subject for each of the 15 indicated conditions were tabulated, averaged, and translated into percentages. Since the number of syllables in the lists is constant the original scores could be used in computing the averages and the translation into percentages made after these averages had been determined for the different quartiles under the varying conditions.

Table I summarizes the results of these computations.

TABLE I

PERCENTAGES LEARNED AT SUCCESSIVE QUARTILE POINTS IN TOTAL LEARNING TIME

Practice Series

Percentile Points	Positions of Lists in Series					Average
	1	7	13	19	25	
25	29.0	36.8	37.8	35.4	38.0	35.4
50	51.8	62.0	64.9	62.4	59.1	60.0
75	72.1	75.3	76.0	77.0	74.7	75.0
100	100.0	100.0	100.0	100.0	100.0	100.0

Continuous Work Series

Percentile Points	Positions of Lists in Series					Average
	1	7	13	19	25	
25	35.5	27.3	32.4	32.7	35.7	32.7
50	60.0	55.1	57.8	58.7	59.3	58.2
75	74.7	72.9	73.1	74.9	74.1	73.9
100	100.0	100.0	100.0	100.0	100.0	100.0

Relearning Series

Percentile Points	Positions of Lists in Work Series					Average
	1	7	13	19	25	
25	39.6	40.4	39.7	43.8	44.0	41.6
50	65.1	66.7	63.4	71.1	72.5	67.8
75	81.9	77.6	76.8	80.8	85.0	80.4
100	100.0	100.0	100.0	100.0	100.0	100.0

In Fig. 1 four typical curves are reproduced: those for the lists in positions 1 and 13 of the practice series, position 13 of the continuous work series, and position 19 of the relearning series. Inspection of the tables indicates that curves plotted

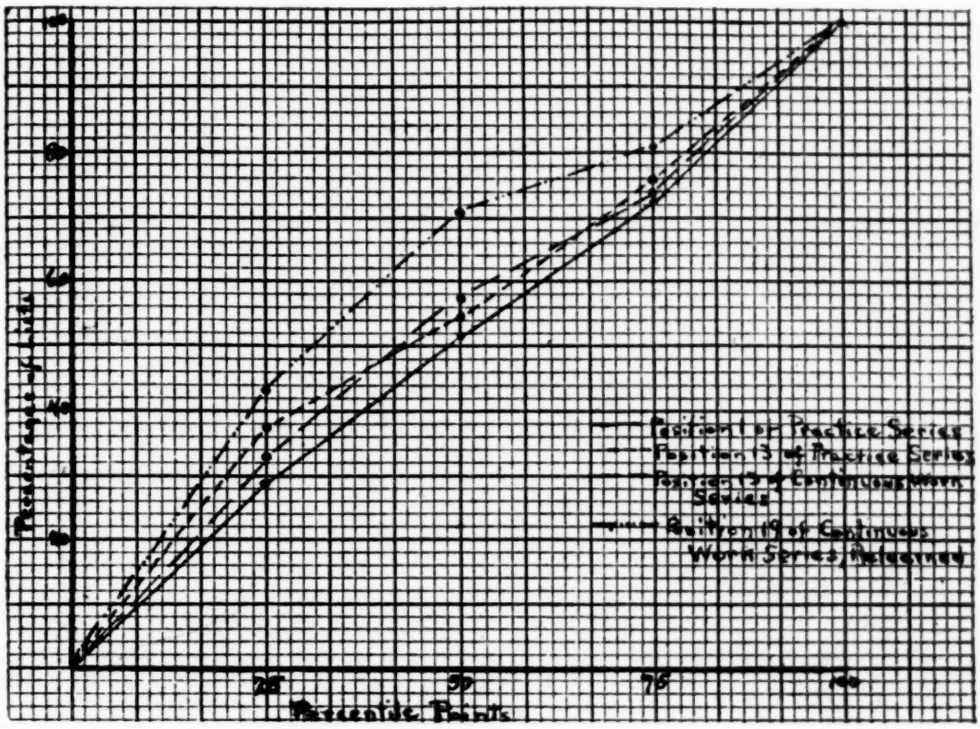


FIG. 1. Percentages learned at successive quartile points in total learning time.

for the ten different positions in the practice series and in the continuous work series will show a remarkably close approximation to coincidence. Two of the curves here reproduced (those for lists in position 13 of each of these two series) are quite typical. The only marked deviation is in the case of the first list of the practice series, in which the earlier repetitions show disproportionately slight returns. This is easily explained as due to the fact that the subjects were not well adjusted to the task so that the earlier presentations were relatively unprofitable. The actual records show that many of the subjects, in learning the first list, required a number of presentations before they could correctly anticipate a single syllable.

The curves for relearning show marked similarity to one another, the one reproduced (that for lists in position 19)

being typical. As might be expected, a significant difference appears when these curves are compared with the curves for the original learning of lists in the practice or continuous work series. The difference is not so marked, however, as might have been anticipated. There is, naturally, a greater advantage in favor of the first quarter and the first half. But this is explained by the fact that we are here measuring not simply the effect of the presentations in the relearning itself but, as well, the persisting effects of the presentations in the original learning. One may well be surprised that the advantage in favor of the earlier quarters is not greater than it actually is. One might reasonably expect that a very considerable percentage of a list might be very quickly relearned, and that only certain particular syllables would require, because of some special difficulty, several additional presentations of the whole list. In such case the resulting curve would show much greater negative acceleration than it actually does. It is to be noted, however, that the results show practically negligible recall scores on the first presentation in relearning, indicating that almost none of the material has been retained at the threshold of recall.

For all the curves the tendency is to ascend sharply to the twenty-five percentile point, less sharply to the fifty percentile point, still less sharply to the seventy-five percentile point, then noticeably more abruptly to the hundred percentile point. This is to say, of course, that the curves are negatively accelerated to the seventy-five percentile point, but show positive acceleration in the final quarter. The first quarter of the learning time gives the largest returns as measured by actual ability to recall; the second quarter is less profitable; the third quarter decidedly less profitable still; while the final quarter stands about on a parity with the second, being inferior to the first but distinctly superior to the third. These tendencies are shown perhaps even more clearly when the percentages learned in successive quarters are tabulated and plotted as in Table II and Fig. 2.

The uniform superiority of the fourth quarter over the third is perhaps the most striking feature of these curves.

TABLE II

PERCENTAGES LEARNED IN SUCCESSIVE QUARTERS OF TOTAL LEARNING TIME

Practice Series

Quarters	Positions of Lists in Series					Average
	1	7	13	19	25	
1	29.0	36.8	37.8	35.4	38.0	35.4
2	22.8	25.2	27.1	27.0	21.1	24.6
3	20.3	13.3	11.1	14.6	15.6	15.0
4	27.9	24.7	24.0	23.0	25.3	25.0

Continuous Work Series

Quarters	Positions of Lists in Series					Average
	1	7	13	19	25	
1	35.5	27.3	32.4	32.7	35.7	32.7
2	24.5	27.8	25.4	26.0	23.5	25.4
3	14.7	17.8	15.3	16.2	14.8	15.8
4	25.3	27.1	26.9	25.1	25.9	26.1

Relearning Series

Quarters	Positions of Lists in Work Series					Average
	1	7	13	19	25	
1	39.6	40.4	39.7	43.8	44.4	41.6
2	25.5	26.3	23.7	27.3	28.1	26.1
3	16.8	10.9	13.4	9.7	12.5	12.6
4	18.1	22.4	23.2	19.2	15.0	19.6

The appearance of this 'final spurt' has been specifically noted only by G. M. Peterson (**2**) in his study of the influence of varying difficulty of material on the form of the memory curve. Kjerstad's curves (**1**) show regular negative acceleration. Robinson and his collaborators emphasize the uniformity of the curves for different types and different amounts of material. This uniformity is in accord with the previous results of Kjerstad and with the results of this study. Inspection of the data presented by Robinson and his associates shows, however, although the investigators do not call attention to it, the occurrence of a 'final spurt' quite like that appearing in Peterson's results and in those here reported.

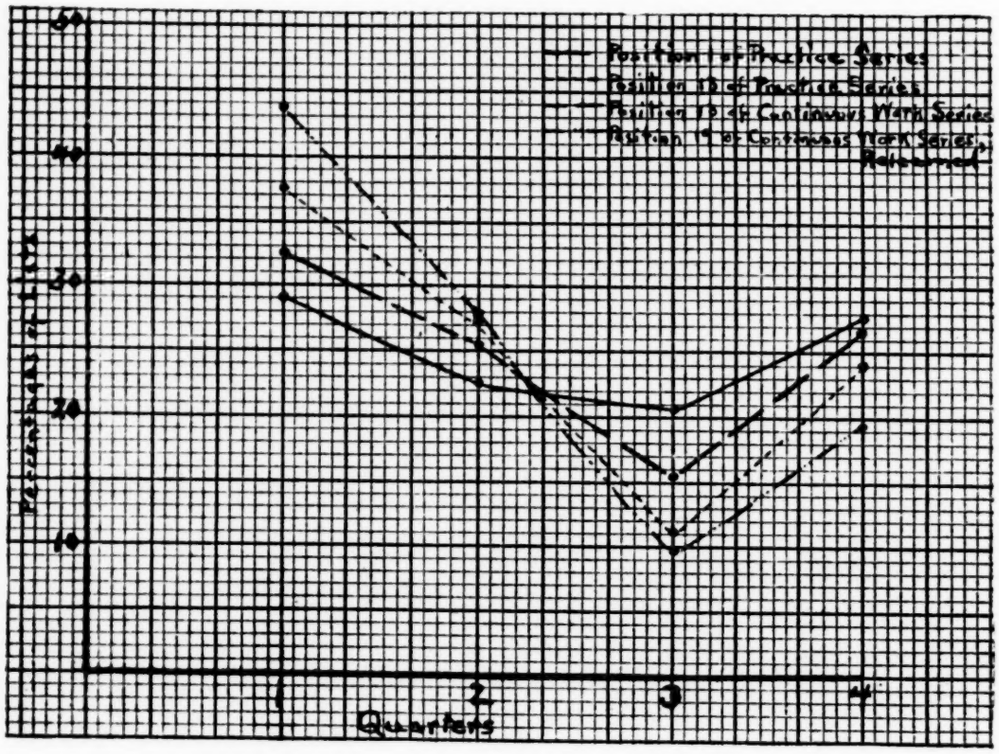


FIG. 2. Percentages learned in successive quarters of total learning time.

We reproduce in full, in Tables III and IV, the tables from the study of Robinson and Darrow (3, p. 239), their Table III representing recalculations made by them, by an improved method, from the results earlier reported by Robinson and Heron. In these tables have been inserted in parentheses figures not occurring in the original tables, representing the percentage of total learning that was accomplished in each successive sixth of the total learning time, since these figures indicate most clearly the characteristic of the curves to which attention is here directed.

It is a very striking fact that with two different kinds of materials, *i.e.*, digits and nonsense syllables, and with four different lengths of lists in the one case and five different lengths of lists in the other, these figures show in every case proportionately greater returns for the final *sixth* in comparison with the preceding sixth, with the single exception of the 4-number list of digits. Here the returns from the final sixth are approximately the same as from the fifth sixth, and both of these are superior to any of the other sixths except

TABLE III

(Robinson and Darrow, Table II)

PERCENTAGES OF LISTS OF DIFFERENT LENGTHS LEARNED IN SELECTED PERCENTAGES (SIXTHS) OF THE TOTAL TIME SPENT IN LEARNING THOSE LISTS

Percentage of Total Learning Time	Length of Lists			
	4 numbers	6 numbers	8 numbers	10 numbers
16.7	19.4 (19.4)	31.8 (31.8)	29.3 (29.3)	30.4 (30.4)
33.3	36.1 (16.7)	44.5 (12.7)	45.9 (16.6)	45.3 (14.9)
50.0	49.8 (13.7)	54.0 (9.5)	55.3 (9.4)	54.6 (9.3)
66.7	63.1 (13.3)	62.7 (8.7)	64.0 (8.7)	64.7 (10.1)
83.3	82.3 (19.2)	71.0 (8.3)	75.8 (11.8)	70.4 (5.7)
100.0	100.0 (18.7)	100.0 (29.0)	100.0 (24.2)	100.0 (29.6)

TABLE IV

(Robinson and Darrow, Table III, representing recalculations from the data of Robinson and Heron)

PERCENTAGES OF LISTS OF DIFFERENT LENGTHS LEARNED IN SELECTED PERCENTAGES (SIXTHS) OF THE TOTAL TIME SPENT IN LEARNING THOSE LISTS

Percentage of Total Learning Time	Length of Lists				
	6 syllables	9 syllables	12 syllables	15 syllables	18 syllables
16.7	23.3 (23.3)	33.9 (33.9)	34.6 (34.6)	30.9 (30.9)	33.1 (33.1)
33.3	45.0 (21.7)	51.2 (17.3)	52.3 (17.7)	52.2 (21.3)	53.2 (20.1)
50.0	60.5 (15.5)	62.0 (10.8)	66.3 (14.0)	62.0 (9.8)	66.3 (13.1)
66.7	72.3 (11.8)	73.6 (11.6)	74.9 (8.6)	74.2 (12.2)	76.3 (10.0)
83.3	83.4 (11.1)	80.8 (7.2)	80.5 (5.6)	80.3 (6.1)	82.8 (6.5)
100.0	100.0 (16.6)	100.0 (19.2)	100.0 (19.5)	100.0 (19.7)	100.0 (17.2)

the first. Thus, even in this one case, there is a 'final spurt,' though it appears a bit earlier than the final sixth. In all the other cases there appear, not consistently negatively accelerated curves, but curves showing a marked 'final spurt' in the last sixth of the total learning time.

Peterson's results indicate clearly enough that the negative acceleration in memory curves is due, at least in part, to the factor of unequal difficulty of the items. If one wishes to speculate further it may be suggested, in terms of the results here reported, that the easier syllables are mastered readily and superiority in actual returns is thus shown in the first

quarter of the learning time. Slightly more difficult syllables are added in the second quarter. In the third quarter, it may be assumed, the more difficult syllables begin to be mastered, but there is a considerable amount of 'partial learning' which does not appear in actual ability to anticipate these syllables. As a result the greater part, or at least a very considerable part, of the effect of the effort in the third quarter appears in actual anticipations in the *final* quarter, thus making the returns in that quarter disproportionately great. (It is of course conceivable that the subject, realizing that he has almost 'got' the list, makes some extra effort toward the end.) The explanation here suggested is an admission that any method of measuring the degree of learning in terms of ability to recall or anticipate actually 'misses' a part of the learning that has actually taken place but the results of which remain below the threshold of recall. That this is the case is a truism. Nor has a method yet been devised to measure this 'partial' or 'subliminal' learning.

The negatively accelerated curve so characteristic of motor learning thus does not appear to be the rule in memorizing. The criteria of progress in the two types of learning are entirely different in character when improvement in motor skill is a matter of increase in speed. Nor is the elimination of errors in motor learning analogous to the bringing of linguistic material to the level of recall. The definition and evaluation of 'errors' is at best arbitrary and beset with difficulties; it is conceivable that under certain conditions the elimination of errors might proceed at a rate quite other than a negatively accelerated one throughout the learning process. In the matter of speed, however, it must be borne in mind that the approach is toward a 'physiological limit' while in memorizing and, as well, in the elimination of errors in motor learning the approach is toward an absolute limit set in terms of one hundred percent correct recall or one hundred percent freedom from the defined errors.

It may be noted, further, that, whatever fractional part of the total learning time be used for the computations on which memory curves are based, the last of these fractional

parts will include the final trial, in which the whole of the list to be learned is recalled. Since this final trial must include at least one addition to the items previously recalled, this fact might seem to weight the returns in favor of the last fractional part and make the 'final spurt' merely an artifact. It may be commented, however, that there does not appear to be any *a priori* reason why the mastery of items in the list should not proceed at a perfectly regular rate, or, for that matter, at a uniformly negatively accelerated, or a uniformly positively accelerated rate. That is to say, while the last trial must always show *at least one* additional successful recall of an item, this does not make it impossible that previous trials might show a successively decreasing number, a successively increasing number, or a uniform number of added items successfully recalled. The successive presentations of a list of twelve syllables might, for example, show returns in the order 4, 3, 2, 2, 1; or in the order 1, 2, 2, 3, 4; or in the order 2, 2, 2, 6; or in the order 6, 2, 2, 2; or in the order 2, 2, 2, 2, 2, 2; or in any one of an indefinite number of orders. Thus would be secured negatively accelerated, positively accelerated, or linear curves with the condition satisfied in each case that the final trial must, because it is final, show measurable improvement. That preceding trials *need* not, and frequently *do* not, show such improvement must be granted. But there seems to be no theoretical reason why, preceding the final trial, there *must be* trials in which *no* measurable progress is shown, and therefore no theoretical reason why the final trial of necessity weights the returns.

Whatever the explanation, the fact seems clear that the curve of memorizing is not universally of the negatively accelerated form, and that at times a 'final spurt' is a noticeable feature of such curves.

References

1. Kjerstad, C. L., The form of the learning curves for memory. *Psychological Monographs*, 1919, **26**, Number 5, Whole Number 116.
2. Peterson, G. M., Negative acceleration with material of varying difficulty. *Journal of Experimental Psychology*, 1928, **11**, 40–44.
3. Robinson, E. S., and Darrow, C. W., The effect of length of lists upon memory for numbers. *American Journal of Psychology*, 1924, **35**, 235–243.

4. ROBINSON, E. S., AND HERON, W. T., Studies in memorization. *Journal of Experimental Psychology*, 1922, 5, 428–448.
5. VINCENT, S., The function of the vibrissæ in the behavior of the white rat. *Behavior Monographs*, 1912, 1, Number 5.
6. WHITE, G. C., *The work curve in memorizing and its relationship to retention.* Abstracts of Theses, The University of Chicago, Humanistic Series, 1927–28, Volume VI, 57–60.

(Manuscript received March 1, 1931)

ANTICIPATION OF CORRECT RESPONSES AS A SOURCE OF ERROR IN THE LEARNING OF SERIAL RESPONSES *

BY FREDERICK HILLIS LUMLEY

Ohio State University

In the many studies of serial learning (nonsense syllable lists and mazes) attempts have been made to describe the elimination of incorrect responses in the different parts of the series. Warden,[1] Husband,[2] and Perkins [3] have shown that human subjects tend to learn the maze in the order: the first part, the last part, and finally the middle part. Ebbinghaus,[4] Meumann,[5] and many others have reported that the same order of learning holds for a series of nonsense syllables. The rat, however, learns the part of the maze nearest the food box first. The data of Hubbert and Lashley [6] and of Husband [2] may be so interpreted. There are other studies in which the results scarcely favor the above generalizations. This is not surprising when the different methods and problems presented to the learner are considered.

From human experience many cases can be cited illustrative of the fact that responses which are due to be made in the near future and which would be correct at that time are anticipated and thus made prematurely. In writing, words are sometimes inverted or unknowingly omitted. In talking, the hurried person may run words together or skip several.

* I wish to thank Professor Samuel Renshaw for his help and criticism during the experimentation and in the preparation of this paper.

[1] C. J. Warden, Primacy and recency factors in cul-de-sac elimination of a stylus maze. *J. Exper. Psychol.*, 1924, **7**, 98–116.

[2] R. W. Husband, A comparison of human adults and rats in maze learning. *J. Comp. Psychol.*, 1929, **9**, 361–77.

[3] N. L. Perkins, Human reactions in a maze of fixed orientation. *Comp. Psychol. Monog.*, 1927, **4**, 92 pp.

[4] H. Ebbinghaus, *Grundzüge der Psychologie*, 1902, **1**, 624–626.

[5] E. Meumann, *Oekonomie und Technik des Gedächtnisses*, 1920.

[6] H. B. Hubbert and K. S. Lashley, Retroactive association and the elimination of errors in the maze. *J. Anim. Beh.*, 1917, **7**, 130–138.

Typists of the author's ability find great difficulty in avoiding transpositions in the eagerness to run off the proper serial response. Furthermore, the tourist finding his way through strange cities either with a map or a vague memory of a previous trip often discovers that he has made the turn too soon.

The problem of the subject learning a maze or a series of nonsense syllables may be looked upon as twofold. First the subject must learn what the nonsense syllables are or just what turns are to be made in the maze. Then the nonsense syllables and turns must be strung together in the correct order. From the learner's standpoint two or more nonsense syllables, one syllable, or even a part of one syllable may be regarded as a unit. It would be even more difficult to define the units in connection with a maze. But for purposes of scoring, arbitrary units such as the turns, alleys, and nonsense syllables may be used.

Since the length of a series is measured in terms of these arbitrary units, we cannot say that two apparently objectively equal series are the same length for the learner. Nor can it be said that one series is twice as long as another. All relationships between the length of the material and the time of learning suffer from this difficulty. This especially hinders the generalization of results obtained from one type of material. Arbitrary units are used in this experiment but with the recognition of their fallibility.

Subjects rarely make an attempt to localize the 'unit' with reference to an exact scale such as could be obtained by counting off the turns or syllables on their fingers. In the experiment to be reported only three of the forty subjects used this method. On the contrary, units are first localized as being in the middle, near the beginning, or near the end. Also certain letters are localized with respect to other letters, and groups are formed. As learning progresses, localization in the series becomes more specific until the series can be given in its right order.

We may assume that the learner has classified five nonsense syllables or turns called A, B, C, D, and E as being 'near the beginning' of the series. In learning a nonsense syllable list he may respond with any syllable of the entire list, but his response will probably be limited to a syllable he knows is 'near the beginning'; for example, D. In the maze

a choice of two directions is offered, one or both of which may be classified as near the beginning. In the former case the learner will usually respond with the correct choice. In the latter case if the direction of the cul-de-sac is the same as the direction of the true path at D, the learner might choose this cul-de-sac. If he chooses D, he *anticipates* [7] a choice which is really three units ahead. Then A is presented to the learner or in the case of the maze he returns from the cul-de-sac and finally takes the correct turn A. Since he is still near the beginning, he will choose one of the other letters, say C. This anticipates the correct response by one unit. Then he is given B and chooses again, this time correctly, C. As the result of classifying the letters in the above way the learner will make fewer errors as he gets nearer the end of the part having the classification 'near the beginning.' This of course will only be true if the learner has a tendency not to try responses which were previously correct. This tendency seems to be present. The only place where the classification 'near the beginning' is not confused to some extent by a subsequent classification such as 'in the middle' is toward the end of the maze. Therefore the learner should eliminate errors fairly easily very near the end of the entire maze or series. This is the recency effect. Further, the working out of the above example will show that the learner should anticipate one unit ahead more frequently than three units ahead. With successive trials, as the localization becomes more exact and the classifications contain fewer units, the relative frequency of anticipating one unit ahead as compared with three units ahead should be increased. For example, if the subject knows that three letters, M, N, and O come after a certain group but does not know in just what order they come, he can anticipate two letters ahead just once and that at the point of commencing the group. He can, however, anticipate one letter ahead twice.

[7] The word *anticipation* refers entirely to the experimenter's classification of the learner's responses. It seems to the experimenter as if the learner anticipates a turn or a nonsense syllable which comes later in the series. Naturally the learner does not sense this anticipation or he would not react as he does. Moreover, since the learner has classified these several turns or syllables as 'near the beginning,' they have equal temporal value and one is as much right from his viewpoint as another.

In brief, the learner should make fewer errors at the very end of the maze than in the middle. He should anticipate syllables or turns which are nearer to the point of choice more often than distant ones. As the maze is learned, the difference between the anticipation of near units and more distant units should be increased. In ordinary maze and syllable sequence-learning we should find a recency effect and a type of anticipation resembling the remote association of Ebbinghaus.[8]

The results of an analysis of published and unpublished studies of errors made in serial learning are given in Table I.

TABLE I

DIFFERENT DEGREES OF ANTICIPATION IN LEARNING MAZES

Results Given as Average Number of Errors

A	B	C	D	Degrees of Anticipation							
Experimenter	Kind of Maze	Number and Kind of Subjects		1	2	3	4	5	6	7	0
Peterson [9]	regular	23	rats	144	127			68			
Peterson [9]	regular	22	rats	100	95						76
Tolman and Honzik [10]	regular	36	rats			10.9		6.5		11.2	3.5
Hubbert and Lashley [6]	Watson C	56	rats		37*	21*					
Husband [2]	Warden T	43	rats			7.1		5.2		7.9	
Husband [2]	W.T. el.	20	humans			5.4		3.0		1.3	
Husband [11]	W.T. el.	20	humans			107		57		25	
Warden [1]	W.T. sty.	60	humans			26.0		10.3		5.1	
Perkins [3]	Eye tube	53	humans	75	60	65	65	47	17		33
Krakoff [12]	Finger	11	humans	9.2	6.2	7.2		5.5			4.2

* Number of trials to eliminate culs-de-sac.

The first column gives the experimenter and the reference. The type of maze used is given in the second column (B). The third (C) and fourth (D) columns give the number and

[8] H. Ebbinghaus, Memory (Ruger translation). New York: Columbia University Press, 1913.

[9] Joseph Peterson, The effect of the length of blind alleys on maze learning. *Behav. Monog.*, 1917, 3, No. 15, 53 pp.

[10] E. C. Tolman and C. H. Honzik, Degrees of hunger, reward, and non-reward, and maze learning in rats. *Univ. Calif. Pub. in Psychol.*, 4, No. 16, 24–56.

[11] R. W. Husband, Human learning on a four-section, elevated finger maze. *J. General Psychol.*, 1928, 1, 15–28.

[12] C. Krakoff. Unpublished report of experiments on learning the Meyer finger mazes, 1930, Ohio State University.

kind of subjects. The remaining columns numbered from 1 to 7 and the zero column give figures for the anticipation of one unit, two units, etc., ahead. The zero column shows that blinds, which did not conform to the true path at any point ahead, or which represented an anticipation of more than seven turns, were present in the maze and gives figures for these. Most of the results are given in averages of the errors for the different kinds of blinds and incorrect responses. To clarify the method of analysis Warden's T maze experiment may be chosen. The maze is so formed that the first cul-de-sac runs in the same direction as the true path three turns ahead. The second cul-de-sac has the same direction as the true path seven turns ahead. The third cul-de-sac, five turns ahead; the fourth, three; etc. Since there are only a few turns of the true path beyond the last cul-de-sac, it was not included in the analysis of this or other mazes. The average of the errors for all the blinds conforming to the true path three turns ahead is given in column three; the average of the errors for all the culs-de-sac conforming to the true path five turns ahead in column five, etc. The other mazes have been analyzed in a similar fashion.

In practically all cases where the true path shortly takes a turn in the same direction as that of the blind there are numerous errors. Perkins [3] notes this result in her study. Furthermore, the culs-de-sac cause less trouble in proportion as the turning of the true path in the same direction is more remote in the forward sequence of turns.

In connection with a modified maze experiment to be reported later [13] an evaluation of the different types of errors revealed this factor of anticipation. Forty subjects learned two 20-letter typewriter keyboard mazes at one sitting and a 40-letter maze a week later. The subject was limited by instruction to a choice of keys *a* to *m* inclusive. He proceeded to press keys until the depression of a certain one gave an auditory signal, and repeated this hunting around until he had pressed twenty keys giving these signals. This procedure

[13] F. H. Lumley. An investigation of the responses made in learning a multiple choice maze. Psychol. Monog. 1931.

was repeated until the subject pressed only the twenty keys in three perfect trials. At any point or choice in the maze the learner had the possibility of making one right and twelve erroneous responses. The letters of the three mazes were: JCLAMILEGALGJMLIKEIA, BALLIFLEDCBAGAGEMBJK, and CHEMICALKGLACIALDELIGIBLEJIMBECILEFLEDGE. Only the first letter in trying to find any right letter was counted. Thus if the learner pressed DF*J*-ABCABCJ*L* in succession in starting on the first maze, only the errors D, A (fourth letter), and A (seventh letter) were recorded and classified. The pressing of A in the above example may be considered an anticipation of the fourth letter in the maze, which is A. The A after the J anticipates the fourth letter by two units and the A after C anticipates the fourth letter by one unit.

Table II and Fig. 1 give the average number of errors

TABLE II

AVERAGE NUMBER OF ERRORS MADE ON THE THREE MAZES, ANTICIPATING THE FIRST, SECOND, THIRD, ETC., LETTERS AHEAD

Maze	1	2	3	4	5	6	7
I	66.7	49.8	43.3	32.9	31.3	31.1	24.7
II	15.1	11.4	9.8	6.8	7.1	8.4	6.4
III	28.6	17.6	22.8	15.5	17.4	9.7	12.4

on the three mazes made by anticipating letters by one unit, two units, three units, etc. These values have been corrected by omitting cases where errors anticipated two or more correct letters instead of one. Had the subject pressed an L immediately after the J which starts the first maze, he would have anticipated both the third letter and the seventh letter of the maze. It is therefore not possible to assign such errors to any one degree of anticipation. This correction and a correction for zero scores have no effect on the general form of the curve, although they smooth the results of the first and second mazes slightly. It is evident that in all three mazes the learners show a distinct tendency to anticipate letters which are nearer to the point of choice than those

which are farther away. Calculation of the averages for a backward association or repetition of previously made correct choices show no special relationship. Such results are given in Table III.

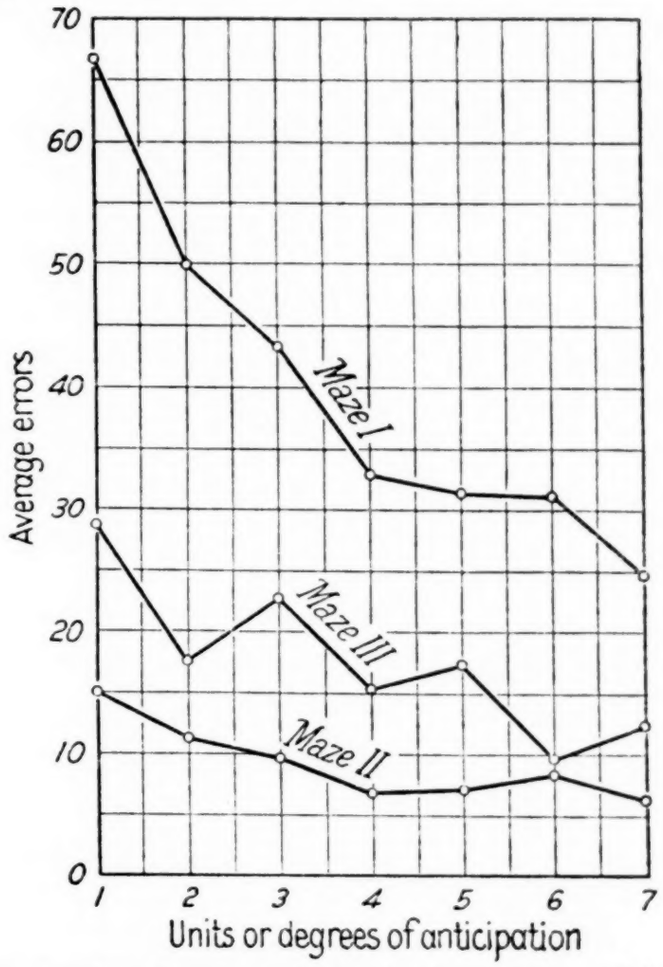


FIG. 1. Average number of errors made on the three mazes, anticipating the first, second, third, etc., letters ahead.

TABLE III

AVERAGE NUMBER OF ERRORS MADE ON THE THREE MAZES IDENTICAL WITH THE FIRST, SECOND, THIRD, ETC., CORRECT LETTERS IN BACKWARD ORDER

Maze	1	2	3	4	5	6	7
I	26.3	29.2	32.8	35.2	38.7	52.4	45.6
II	11.6	10.0	10.4	7.2	16.6	9.5	10.4
III	13.6	12.7	17.6	16.0	16.4	14.6	14.5

Table IV shows the data of the first maze treated in a different fashion to provide a check on the relationship made evident in Table II.

TABLE IV

FIRST MAZE. ERRORS ARE GIVEN AS THE AVERAGE OF THE PERCENTS OF THE TOTAL ERRORS MADE ON THE DIFFERENT LETTERS, ANTICIPATING THE FIRST, SECOND, ETC., LETTERS AHEAD

Treatment	1	2	3	4	5	6	7
All scores	24.5	15.8	11.0	13.4	14.4	9.2	10.0
Repetitions and zero scores removed	23.3	15.7	14.9	12.8	9.6	9.3	7.4
Habit interference and above scores removed	23.3	13.6	11.7	11.1	10.5	9.3	5.9
Ebbinghaus. Saving of work in percent of original learning time (results doubled). (1)	21.6	14.0	11.6				6.6
(2)	23.8	15.0	12.2				6.6

The middle part of the maze was learned last and consequently a massing of errors occurred there. The errors for each degree of anticipation are accordingly given in terms of the percent of the total errors made at each point in the maze or series. This has the effect of making the number of errors made in locating every correct letter in the maze the same, and therefore compensates for primacy and recency effects. The first row in Table IV shows the data in this form. The second row gives the data with the additional correction mentioned in connection with Table II, and the third row shows the results minus the cases of habit interference caused by having letters repeated in the maze. The results confirm completely the tendency found for all the mazes by the other method.

Since it might be expected that the effect of this anticipation would become apparent in relearning experiments in the form of remote association the results of Ebbinghaus' [8] two experiments are also given in Table IV for comparison. The values have been doubled to make the general relationship and similar slope of the curves stand out. The interesting question of what would happen to remote association if anticipatory responses were eliminated remains undecided, although Cason [14] has suggested, and found to some extent, that proper control of the learning process prevents the appearance of remote associations in relearning tests.

The errors made on maze two were divided into four

[14] H. Cason, Specific serial learning: a study of remote forward association. *J. Exper. Psychol.*, 1926, 9, 299–324.

parts according to the stage of mastery of the maze. The errors made on the first trial were put into part one and the rest of the errors were distributed as equally as possible. The sum of the last three (far) anticipations (columns 5, 6, and 7 in Table II) was divided by the sum of the first three (near) anticipations for each part. The ratio is 2.223 for the first part. For the second, third, and fourth parts the ratios are .577, .459, and .316 respectively. As the subject learns the second maze, he makes more near anticipations and fewer far anticipations. This conforms to the previous statement that the difference between the anticipation of near units and more distant units should be increased as the maze is learned.

A slight recency effect is observable in the learning of each of the three mazes.

Anticipation may also prove to be a factor in ant life if my analysis of the published results of Dr. Schneirla's [15] maze learning by ants is correct. Table V presents the analysis of the data. Although the number of cases is small and a maze with more blind alleys is needed to demonstrate the

TABLE V

AVERAGE NUMBER OF ERRORS MADE BY SCHNEIRLA'S ANTS ON CULS-DE-SAC HAVING THE SAME DIRECTION AS THE TRUE PATH A GIVEN NUMBER OF TURNS AHEAD

Maze	Ant	Degree of Anticipation					
		To Food			Return to Nest		
		3	4	5	2	6	0
Btl	o	14	3	2	9		1
	A	30	2	8	8.5		2
	3	33	1	6	15.5		4
Av.		22.3	2	5.3	11.0		2.3
Bbl	dot	45		9		1.5	0.5
	d	43		7		4	6
	z	47.6		2		5	0.5
Av.		45.2		6		8.5	2.3

[15] T. C. Schneirla, Learning and orientation in ants. *Comp. Psychol. Monog.*, 1929, 6, No. 30, 143 pp.

presence or absence of this factor, the results show that blind alleys having a direction similar to that which the true path takes a little further on are entered more often. Dr. Schneirla, however, describes the action of these ants in terms of centrifugal swing, and it is possible that the apparent presence of the anticipatory factor may be only a result of this.

Conclusion.—The problem of the selection of the true path and the elimination of errors has been a troublesome one in both animal and human learning experiments. As an approach to the solution of this problem various attempts have been made to account for the types of errors made. Frequency, recency, goal-orientation, length of alleys, backward elimination, etc., have been suggested as factors. This paper tries to show that, along with or in spite of these factors, errors suggesting an anticipation of responses which would be correct further along in the series are made. The remote association results obtained in relearning experiments may be attributed to the operation of this factor during the original learning. No evidence of a similar factor for backward association was found.

In conclusion it must be admitted that these data are based upon arbitrarily chosen units in the hope that these will correspond somewhat to the learner's analysis of the series. Observations and reports of the learner show that from his standpoint the units and groupings are much changed during the course of learning. A more accurate study of anticipation can only be made with constant reference to these changing units.

Summary.—On the basis of the assumption that one of the problems of a subject learning a series of responses or a serial response is the localization of the units composing the series, three effects should be observable in the subject's record:

1. The terminal end of the series should be learned more quickly than the middle.

2. The subject should show an anticipation of units further forward in the series in inverse proportion to the distance of these units from the point of choice.

3. As learning progresses, the ratio of the far anticipations to the near anticipations should be decreased.

The results of the experiment reported in this paper confirm these three conclusions. The data of other investigators when analyzed in a similar way support the second statement.

(Manuscript received February 26, 1931)

CARD SORTING AS A MEASURE OF LEARNING AND SERIAL ACTION

BY MILES A. TINKER, ARTHUR J. IMM AND CHESTER A. SWANSON
University of Minnesota

The wide use of card sorting in both learning and serial reaction experiments indicates its popularity and its availability.[1] Although a large number of investigators have employed some form of card sorting for one purpose or another, very few have attempted an analysis of performance in the various kinds of sorting. The present study was planned to determine the reliability of performance, the relation between performances, and progress in learning for various types of card sorting.

Apparatus and Procedure

In this experiment the deck consisted of 60 cards, 15 in each suit. The cards were dealt or sorted into four shallow boxes about 6 by 8 inches. In all parts of the experiment the boxes remained in a constant position, each forming a quadrant of an oblong area on a table in front of the subject. Each box carried a section of a card to indicate one of the suits.

Twenty-five university students served as subjects in the main experiment and 20 in the supplementary study. There were about an equal number of men and women in each group.

In the main experiment the cards were distributed in five different ways (W, I, II, III, IV); in the supplementary experiment in three ways (W, II, IV). The first method of distributing (W) the cards was designed as a warming up exercise and to measure speed of response where choice is at a minimum, or absent. Holding the deck face down, the cards were distributed, one at a time and without regard to suit, into the four boxes, rotating in a clockwise manner. In I the deck was held face up and the cards sorted into their respective boxes according to suit. Type II was the same as I except that the pack was held face down which necessitated the turning over of each card before a choice of its compartment was made. In III the 'one-behind method'[2] of sorting was employed with the deck held face up. Here the top

[1] For references on card sorting as a measure of serial action see C. F. Hansen, Serial action as a basic measure of motor capacity, *Psychol. Monog.*, 1922, 21, 320–382; and R. W. Husband and W. R. Miles, On sorting packs of sixty cards with form and color as variables in two to six kinds, *J. Appl. Psychol.*, 1927, 11, 465–482. For use in learning experiments see W. H. Pyle, *The psychology of learning*, Warwick and York, Baltimore, 1928.

[2] See Husband and Miles, *op. cit.*

card was dealt face up into the upper left box. Then the second card was thrown into the box carrying the same suit designation as the first card dealt. Similarly card two indicated where card three was to go, etc. Sorting procedure IV was the same as III except that the deck was held face down.

Following a practice period, the main experiment consisted of eight sittings, at each of which one trial in every procedure was completed. One to three days intervened between sittings. The warming-up (W) was always done first but order of sorting in the other procedures was systematically varied to equate fatigue, practice, interference,[3] etc.

In the supplementary experiment, the cards were distributed by procedures W, II and IV and for four trials only. The four trials in each procedure were completed at a single sitting during which the order of sorting was systematically varied.

All subjects were instructed to correct all errors they noticed. The score was number of seconds taken to distribute the deck of 60 cards. No errors were recorded since they are infrequent and non-diagnostic in card sorting.

Results and Discussion

The reliabilities of the average scores in card sorting were computed by correlating sum of scores on even trials with sum on odd trials and applying Brown's 'prophecy' formula. The reliability coefficients for both groups of subjects follow:

W. Main group	$r = .82$,	$r_x = .90$
I, Main group	$r = .88$,	$r_x = .94$
II, Main group	$r = .91$,	$r_x = .95$
III, Main group	$r = .83$,	$r_x = .91$
IV, Main group	$r = .82$,	$r_x = .90$
W, Supplementary group	$r = .96$,	$r_x = .98$
II, Supplementary group	$r = .87$,	$r_x = .93$
IV, Supplementary group	$r = .81$,	$r_x = .90$

The size of these coefficients demonstrate that card sorting, as done in this experiment, is highly consistent from trial to trial and from day to day. The raw coefficients range from .81 to .96 and when corrected, from .90 to .98. The coefficients are quite comparable for the two groups except in procedure W where the reliability is higher in the supplementary group where all trials were completed at a single sitting.

Table 1 gives the average number of seconds taken to distribute 60 cards in each type of sorting for both the main group and the supplementary group (procedures W, II and

[3] Due probably to the nature of the sorting and to the constant position of the sorting compartments very little habit interference was apparent during the experiment.

IV only). The latter group was tested to check trends discovered in the main group. As one would expect, dealing (W) is done more quickly than any other kind of sorting and its variability is very small. Sorting in suits with the deck

TABLE 1

AVERAGES AND STANDARD DEVIATIONS IN SECONDS FOR DISTRIBUTING CARDS IN VARIOUS KINDS OF SORTING

Kind of Sorting	Main Group, $N = 25$		Suppl. Group, $N = 20$	
	Av.	$\sigma_{dis.}$	Av.	$\sigma_{dis.}$
W (Dealing)	25.1	2.0	22.8	1.9
I (In suits, face up)	37.3	5.0		
II (In suits, face down)	64.7	6.6	64.7	6.8
III (One-behind, face up)	61.9	11.4		
IV (One-behind, face down)	61.0	9.6	69.1	10.9

held face up (I) is next fastest. Turning a card over prior to choice of the box into which it must go (II) apparently slows up the response considerably for sorting in II is much slower than in I; 64.7 vs. 37.3 sec. In procedures III and IV however, it makes no difference whether the deck is held face up or down. The respective averages are 61.9 and 61.0 sec. The reason for the above differences becomes apparent on comparing methods of sorting. In I and II the subject must look at the face of each card before choosing the box into which it goes. Turning a card over (in II) delays the choice. In types III and IV, on the contrary, choice is made by viewing the previously delt card as it lies in the box. Although choice is more complex (difficult) in the early trials of III and IV, greater improvement occurs for a given number of trials than in either I or II. The total averages of III and IV are slightly less than for II, because the last four trials in both III and IV are sorted much more rapidly than in II. Variability in improvement will be discussed further under learning.

Averages for the supplementary group reveal the same trends discovered in the main group except for the relatively longer time of response in procedure IV. As only four trials

were completed for the supplementary group the longer times in the early trials of IV are reflected to a greater extent in the average than in the corresponding average of the main group.

Both the absolute and the relative variabilities show the same trends in both groups. Variability increases in the order W, II, IV. In the main group the relative variability of I is greater than II, and of III is greater than IV.

From the basic data for learning to sort by the various methods Fig. 1 was drawn. Marked differences are revealed in progress with practice. The learning curves indicate that

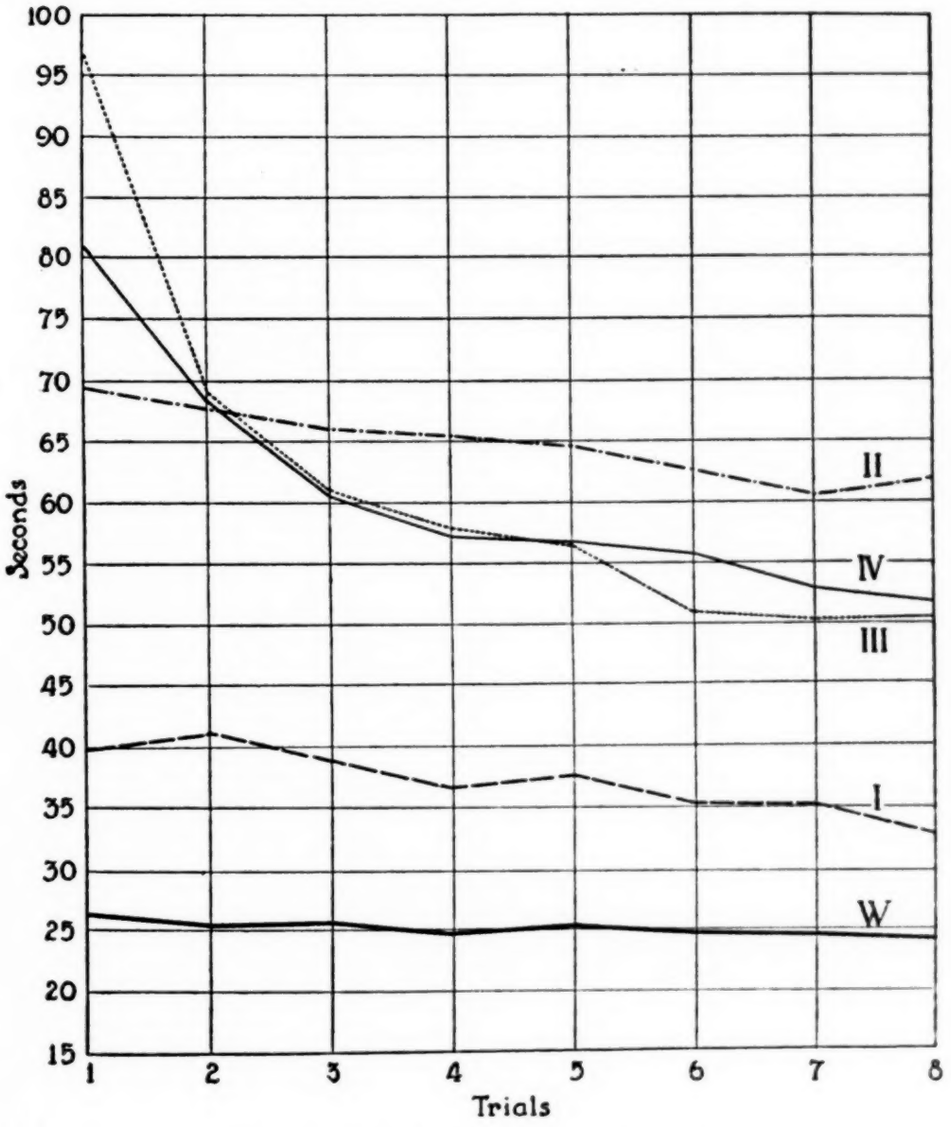


FIG. 1. Learning curves in card sorting.

(1) practice in procedure W produces only very slight improvement; (2) practice in I and II results in small and approximately equal amounts of improvement, but on markedly different levels; (3) practice in III and IV is accompanied by a large amount of improvement which yields the 'typical,' negatively accelerated learning curve. These two curves follow each other very closely except on the initial trial where the greater efficiency of IV is due to its following III with all subjects during the first sitting. The relatively greater improvement in III and IV over I and II is undoubtedly influenced by the fact that the sorting movement with practice, becomes one swift, continuous coördination with the vision directed entirely to the trays. In I, on the other hand, choice must be made before dealing the card, and in II occurs the complicating turn of each card before choice. Both III and IV permit more overlapping of mental (choice) and motor sequences than either I or II.

For a few subjects in both groups (three subjects altogether) the practice was continued through 20 trials. Slight or no improvement occurred from the 8th to the 20th trial: W, 19 sec. on the 8th trial to 19 sec. on the 20th trial; II, 55 sec. to 51 sec.; III, 50 sec. to 50 sec.; IV, 49 sec. to 48 sec. Therefore the first eight trials yield a representative picture of progress with practice in these types of card sorting.

From the above one may conclude that variability and speed of response as well as progress in learning is affected by changes in either motor sequences or complexity of choice in sorting cards into compartments whose position and suit remain constant.

Correlations between scores in the different kinds of sorting are given below:

Main Group

Dealing (W) and sorting into suits, face up (I)	$r = .36 \pm .12$
Dealing (W) and sorting into suits, face down (II)	$r = .49 \pm .10$
Dealing (W) and one-behind, face up (III)	$r = .33 \pm .12$
Dealing (W) and one-behind, face down (IV)	$r = .45 \pm .11$
I and II	$r = .67 \pm .07$
I and III	$r = .74 \pm .06$
I and IV	$r = .74 \pm .06$
II and III	$r = .53 \pm .10$
II and IV	$r = .52 \pm .10$
III and IV	$r = .89 \pm .03$

Supplementary Group

W and II..... $r = .57 \pm .10$
W and IV..... $r = .30 \pm .14$
II and IV..... $r = .48 \pm .12$

In the main group simple dealing (W) and sorting into suits with the face down (II) correlates low with all other forms of sorting; *r*'s range from .33 to .53. Distributing into suits with face up (I), however, correlates fairly high with all except dealing; *r*'s range from .67 to .74. The highest intercorrelation occurs between the two forms of the 'one-behind method'; $r = .89$. Correlations in the supplementary group reveal the same trends as those discussed above.

Consideration of differential response in sorting and the intercorrelation coefficients indicate that the following are among the factors which favor equivalent scores in card sorting: (1) smooth coördination in motor sequences (I, III, IV); and (2) overlapping of motor sequences and processes of choosing (I, III, IV). Similarly, factors which tend to lessen correspondence in performance include: (1) variation in complexity of choice; and (2) interrupting the continuity of motor sequences (*i.e.*, turning a card over to see the suit before choice).

Summary and Conclusions

1. Speed measures for distributing cards in five different methods by 25 subjects and in three methods by 20 additional subjects were obtained. The first group had eight trials, the second, four.

2. All methods of card sorting were found to be highly reliable, the corrected coefficients ranging from .90 to .98.

3. Variability and speed of response as well as progress in learning is affected by changes in either motor sequences or complexity of choice in sorting cards into compartments whose position and suit remain constant.

4. Factors which favor equivalent scores in different forms of card sorting include (1) smooth coördination in motor sequences, and (2) overlapping of motor sequences with processes of choosing. Among factors tending to lessen equivalence of scores are: (1) variation in complexity of choice; and (2) interruptions of continuous motor sequences.

(Manuscript received February 7, 1931)

HYPNOTIC SUGGESTION AND THE CONDITIONED REFLEX[1]

BY V. E. FISHER
New York University

In performing the experiment reported here, the writer had no new hypothesis to test, no theory to prove or disprove. The investigation was purely of an exploratory nature, the results have only a tentative and qualitative significance, and the interpretation of the findings will be left largely to the reader. The general purpose of the experiment was to determine in what manner, if any, a conditioned reflex response might be influenced by hypnotic suggestion.

Subjects, Apparatus, Procedure

Five volunteer female subjects, students in good standing at New York University, were employed. The patellar reflex (knee-jerk) was selected because of its accessibility to experimental methods and its comparatively involuntary nature. An ordinary door-bell was used to supply the conditioned stimulus, so altered and wired as to give a single clang upon the pressing of a button. The unconditioned stimulus was obtained by means of a simple mechanical device holding a small rubber tipped hammer which could be readily adjusted to different positions. Tension for operating the hammer was secured by the use of elastic bands, thereby reducing noise to a minimum. Inasmuch as the laboratory was practically sound proof, the only appreciable sounds were the clang of the bell, the light thud of the hammer against the subject's leg, a light hum of the kymograph, and the click of the markers on the kymograph drum. It should be added, however, that besides the clang proper of the bell, there was a duller sound such as is produced by striking a piece of metal against a second piece lying on soft wood or earth. During the final tests, kymographic records were made of the conditioned stimulus (the bell) and of the subject's response (movement of the leg).

Inasmuch as the five subjects differed considerably in degree of hypnotizability, a few words concerning them is necessary. It was not possible to obtain any signs of the deeper states of hypnosis in three of the subjects. The fourth subject readily carried out post-hypnotic suggestions of the motor-reaction type and was amnesic for the hypnotic period; but such phenomena as post-hypnotic anæsthesias (negative hallucinations) could not be induced, or at least only partially. The fifth subject (M.S.) was an ideal experimental subject for study in hypnosis. To all appearances post-hypnotic amnesia was complete in her; such phenomena as post-hypnotic anæs-

[1] This experiment was made during the spring of 1928. It was reported at the 1930 meeting of the New York Branch of the American Psychological Association.

thesias (auditory, visual, tactual), paralyses, contractures, 'compulsions,' automatic writing, etc., could be easily and quickly induced. Because of considerations for space, expense, etc., together with the fact that M.S. was the only satisfactory subject of the five for this type of experiment, only her records are here reproduced.

During the period of conditioning, the subject was required to sit on the edge of a small laboratory table with legs hanging freely over the side. Before and during the course of the experiment she was given no instructions other than to take her accustomed position on the table, to refrain from discussing the experiment with other individuals, including the experimenter, and to avoid trying to anticipate the purpose or significance of the experiment. All of the subjects appeared to coöperate in the most willing manner.

The usual method for conditioning a response was employed, without, however, adopting a fixed time interval between combined stimulations or between the applications of the conditioned and unconditioned stimuli. About eighteen hundred combined stimulations were given to each subject, two hundred stimulations being given at each practice-period, three practice-periods a week. It was necessary to take the final readings at the end of this time on account of the final examinations and the closing of the school year.

The conditioning appeared to be fairly complete in M.S. at the time the final records were made. The degree of conditioning obtained in the remaining four subjects differed appreciably, and although no careful observations were made concerning this variation, the writer is of the opinion that the conditioning tended to progress more or less in proportion to the hypnotizability of the subject. The response was conditioned in M.S. while she was in the hypnotized state.

Results for Subject M.S.

The accompanying graphs, Fig. 1, give the results obtained with this subject under the ten following conditions:

1. Hypnotized, blindfolded, no control.
2. Waking state, blindfolded, no control.
3. Hypnotized, cancelling letters behind a screen.
4. Waking state, cancelling letters behind a screen.
5. Hypnotized, blindfolded, voluntary effort to inhibit response by trying to hold foot still.
6. Waking state, blindfolded, voluntary effort to inhibit response by trying to hold foot still.
7. Waking state, blindfolded, following post-hypnotic suggestion that she would not hear the bell.
8. Waking state, blindfolded, voluntary effort to ignore sound of bell—no suggestions given.
9. Waking state, blindfolded, following post-hypnotic suggestion that it would be impossible for her to keep her attention from the bell, that her attention would be 'rivited to the bell.'

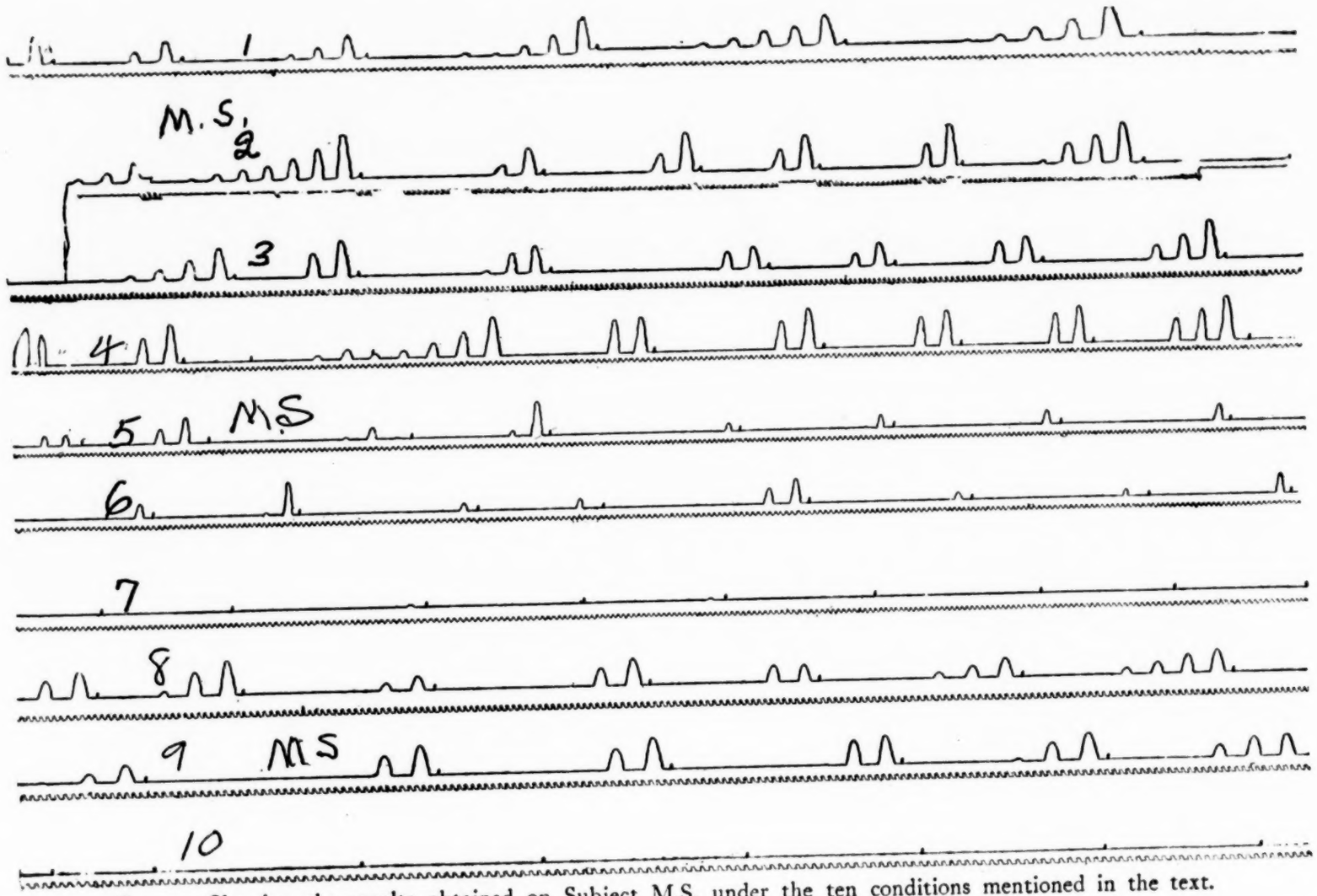


FIG. 1. Showing the results obtained on Subject M.S. under the ten conditions mentioned in the text.

10. Waking state, blindfolded, following post-hypnotic suggestion of complete auditory anæsthesia for all sounds.

The numbers from '1' to '10' designating these different conditions correspond to the respective numbers on the records and the tests were made in the order given above. The more regular, lower line in each record is a time-line marking fifths of a second. It has little or no significance in the interpretation of the results. The more irregular upper line contains records of the conditioned stimulus (the bell) and of the subjects response (the knee-jerk). The small sharp break in the line signifies the occurrence of the stimulus, the sequence of larger breaks, instead of only one, was due to the swinging back and forth of the leg. The height of the first break—immediately following the break produced by the stimulus—is approximately proportionate to the size of the arc described by the subject's foot. The records are to be read from right to left.

Following alternate conditions, the subject was given fifty combined stimulations, thus, conditions '1' and '2'—50 combined stimulations—conditions '3' and '4'—50 combined stimulations, etc.

Interpretation, Conclusions

Under conditions '1,' '2,' '3,' '4,' '8' and '9' there appears to have been but little interference with the response; or perhaps I should say that the degree of interference (inhibition, reinforcement) if any, was of such degree as could be detected and measured only by careful quantitative methods. Under conditions '5,' '6,' '7' and '10,' on the other hand—regarding the responses under conditions '1' and '2' as a sort of norm or standard—there was a very apparent inhibition. Under conditions '5' and '6' the subject was given no suggestion to the effect that she would hold her foot still; rather she was commanded during the period of actual stimulation to hold her foot still. Under condition '5' (hypnotized state) after she had been commanded several times to hold her foot still, she complained that her leg was hurting her and appeared to be on the verge of weeping.

Her leg had become very rigid (hypnotic catalepsy). It is interesting to note that there appears to have been very little difference in the degree of inhibition under these two conditions ('5' and '6').

A glance at the records shows that under condition '7' there was a slight response (overt) twice; that for the other stimulations there was no response. I took occasion to ask the subject following this reading if she had heard anything. She replied that the only thing she had heard was a slight clicking sound, which was later identified as a sound sometimes produced by a loose clamp when the bell was sounded. It will be recalled that complete auditory anæsthesia was not suggested for condition '7,' but only anæsthesia for the bell itself.

Under condition '10,' a post-hypnotic state for which complete auditory anæsthesia had been suggested, no detectable response whatever was made. Following this reading, while I was removing the blindfold, the subject said, "But I thought you were going to take a reading." When I told her that I had she grew somewhat perturbed, declaring that she must be abnormal in some way if she could sit within two or three feet of the bell and never hear it.

A few words are in order concerning the results obtained with the other four subjects. Only in one was I able to secure a condition at all approaching post-hypnotic auditory anæsthesia. Following a post-hypnotic suggestion to this subject that she would hear absolutely nothing (condition '10') she reported that the only thing she had heard was the singing of birds, and then immediately became puzzled for she knew quite well that birds of that sort never gather around Washington Square College. She had made an illusory perceptual response to the bell. She responded part of the time to the stimulations, part of the time she did not. Otherwise her results were very comparable to those of M.S., except for condition '7' which was omitted. For the remaining three subjects the results varied a good deal. I am of the opinion that the conditioning in these three subjects was not sufficiently complete to give their results any significance.

I shall refrain from generalizing on the basis of results obtained from a single subject. I think I may legitimately point out, however, that as far as they go the results point to a dependence of a conditioned response, even of a fairly pure reflex, upon conscious perception of the stimulus.

The experiment suggests certain possible interesting and fruitful lines of investigation. Could a reflex response be conditioned to a sound stimulus in a subject during post-hypnotic auditory anæsthesia? Does conditioning of a simple response tend to take place more rapidly under hypnosis than in the waking state? Could an unconditioned reflex response be altered or eliminated (inhibited) by means of post-hypnotic suggestions of anæsthesia?

(Manuscript received February 3, 1931)

A NEW APPARATUS FOR MEASURING CHOICE AND DECISION

BY MAX F. HAUSMANN

Department of Psychology, Johns Hopkins University

With the purpose of devising an experimental situation for a psychobiological choice and decision reaction, we constructed an apparatus in the Psychobiology Laboratory of the Phipps Clinic during the first months of 1930.

The apparatus is essentially a multiple choice apparatus. The subject responds to a number of light and sound stimuli by pressing the corresponding keys. The stimuli are presented automatically and the records of the experiments are taken automatically also. In order to give the situation a more realistic setting, the experimenter informs the subject that the signals correspond to some of the hospital call bells and it is therefore important to relay them to the main switchboard promptly and correctly.

The choice principle is that in the case of a number of signals, the subject is told that he may react in two ways, e.g., he can either choose to answer by operating a key located at a certain distance (several yards, other room, etc.), or he can choose to respond by operating a key immediately before him, but charged with electricity: the choice situation is therefore one of deciding which is the greater evil, 'getting up and taking the trouble of walking to the switch and back,' or 'enduring an electric shock.'

These two alternatives are only indicative of what can be done with the apparatus. Each experimenter can devise for himself with the aid of the apparatus, choice situations of a similar nature, by varying systematically the nature and intensity of rewards and punishments. For instance, a piece of candy can be used to tempt the subject to walk the longer distance. In another case a differentiation could be expected between walking to the switch when a small distance is called

for, and taking a chance on the electric shock when the distance to walk becomes too much trouble. Similarly, variations can be expected at different hours of the day, or after a certain duration of the experiment.

The apparatus is at the present time unfortunately still in a somewhat crude state, but a heavier model will be constructed soon. This model is expected to be better adapted for the strain of prolonged experimentation.

The problem put before us, early in 1930, was to 'devise a multiple choice apparatus (*a*) with entirely automatic control; (*b*) capable of operating for many hours in succession, and (*c*) at as low expense as possible.' In view of the somewhat novel purpose we intended to use the apparatus for, a few modifications from the traditional set-up had to be made.

The apparatus as now set up consists of five essential parts, these being respectively: (*a*) the actual signals (light and sound) and the response switches; (*b*) the timer or controller, *i.e.*, the apparatus initiating the operation of the signals; (*c*) the relay system; (*d*) the recording system; (*e*) the batteries.

(*a*). The number of signals is now five: 4 lights and 1 buzzer signal. Three of the lights are mounted together on a board about $2\frac{1}{2}$ feet long. The fourth light and the buzzer can be placed at various places in the room. Ordinary flashlight bulbs for $2\frac{1}{2}$ volts are used; the buzzer is a standard 3-volt buzzer. The response switches are ordinary break switches as used for lighting and similar purposes.

The arrangement is such that two signals, A and B, can be responded to in two ways: In the case of A, the subject has the choice between walking about ten feet (in our case; of course, any distance can be arranged for) in order to pull a response switch; and pulling a switch placed immediately before him. This latter switch is, however, under an electric charge, and the subject receives a moderately severe electric shock when he decides to operate this switch.

In the case of signal B, the choice is the same, except that the subject does not receive an electric shock, but is punished 'morally' by receiving a demerit on his score.

This arrangement seems to present a psychobiological choice and decision situation; it is, as far as we are aware, a new principle in multiple choice devices and allows a great number of modifications. The underlying idea is, however, always the same: "which of two punishments (or rewards) do you prefer at a given moment." This principle should allow to differentiate rather accurately between concomitant desires at different moments.

The shock mentioned is administered by an induction coil, one pole of which is grounded, the other being connected to the switch in such a way as to make pulling the switch impossible without receiving the shock. The intensity of the shock can be varied. The primary circuit of the induction coil is electrically separated from all the other circuits; it is operated through a relay and a separate battery.

Each signal consists of a 'stimulus' (light or sound) and a 'response' key (switch) placed immediately beside it. A learning process for the association of stimulus and response is therefore hardly involved.

(*b*). The Controller: consists of a series of three wooden cylinders about one foot long (rolling pins!) rotated at a slow rate; the fastest one makes one complete revolution in fifteen seconds; the slowest one rotates once in about four minutes. The power is furnished by a small 110 volts DC electric motor, the speed being stepped down essentially by a worm gear of a ratio of 1 : 50 and a series of belt transmissions.

Five pairs of narrow phosphor bronze springs press against each of the wooden cylinders as they rotate. On the cylinders there are fastened horizontally a certain number of small, short copper sheet pieces in such a way as to make contact between the two springs of one pair as they pass over them on the revolution. The number of contacts made for each of the spring pairs per revolution depends therefore on the number of copper sheet pieces encountered. The arrangement is the same for the three cylinders, and spring pair I of the first cylinder is hooked up in series with spring pair I of the second cylinder and of the third one:

It is therefore evident that an electric current can pass through the system only if the three spring pairs happen to meet one of their respective copper piece all at the same time.

This arrangement has the effect of decreasing the number of current impulses per time unit very considerably and of producing a wide variety of the sequence of the signals. The same order does not recur until after a considerable lapse of time.

(*c*). The Relay System: consists of five relays type Western Electric A-26 (single winding; two break contacts). One relay (R1, Schematic I and II) is used for each signal. The relays have three functions, viz.: (1) They transform the momentary impulse received from the controller springs into a permanently closed circuit; (2) they open this circuit as soon as the response switch is operated, *i.e.*, they turn the signal off; (3) they signal these operations to the recording system (through contacts A_2 and B_2). Relays of the same type are used for the operation of the choice mechanism of the signals A and B (R2, Schematic II). They are operated by the distant response switch for A and B, and interrupt their signal currents (by breaking the contacts A_3 and B_3). At the same time they send an impulse to the recording system (through A_4 and B_4), thus recording that the subject has chosen to walk to the distant switch, and that he has not responded by touching the charged switch.

The relay operating signal A is operating at the same time an auxiliary relay (R3), which in turn closes the induction coil circuit (through contact A_5 and B_5), and thus puts switch I under electric tension. The auxiliary relay is released when the signal current is cut off.

(*d*) The Recording System consists of electric markers.

(*e*) Ordinary Eveready cells have been used for our purpose. We have grouped them in four batteries:

(1) One battery is used to operate the signals and the relays (Bt-1).

(2) One battery to operate the electric markers (Bt-2).

(3) One battery to operate the 'choice' relays (Bt-3 on Schematic II).

(4) One battery to operate the induction coil (Bt-4, on Schematic II). A bipolar master switch, not indicated on the two drawings, allows to cut off batteries 1 and 2; this automatically makes the operation of the other batteries impossible.

Summary

A multiple choice device is described, deviating from the conventional types in a number of respects:

1. The device gives the signal entirely automatically over any period of time.

2. The signals do not appear in the same sequence for a very prolonged period of time.

3. The recording is done automatically.

4. A psychobiological choice and decision situation is presented to the subject.

5. The cost of the apparatus is moderate in the construction as well as in the up-keep.

APPENDIX

Schematics Showing Electric Hook-up

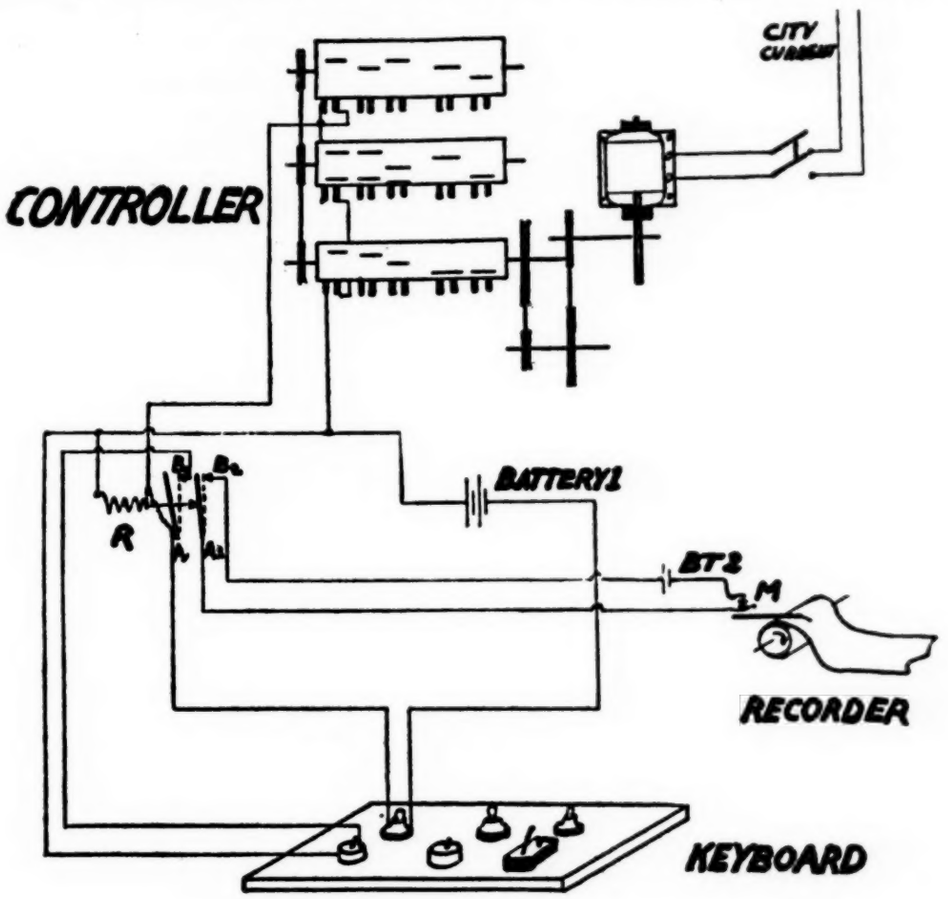
Schematic I

1. *Rest.*—A very small resting current (appr. 0.02 amp.) flows in the system Bat–R–B–Bulb–Bat. The current is too weak to light the bulb, but strong enough to keep the relay R magnetized—therefore the contacts A_1, B_1 and A_2, B_2 are open. The *controller* turns at a very slow rate; at a given moment the two springs (*i.e.*, three spring pairs) corresponding to the signal represented in the drawing are short circuited through the metal plates on the controller. This initiates stage shown on opposite page.

2. *Operation.*—Through the controller spring system the winding of R is short circuited, becomes currentless and releases its contact springs. Thus contacts A_1, B_1 and A_2, B_2 are closed: this short-circuits the relay permanently, eliminates the high resistance of the relay and allows a current strong enough to light the bulb to pass through the system. At the same time, contacts A_2, B_2 close a current operating

the marker. This stage represents the *stimulus*. The *response to the stimulus* or *reaction* leads to

3. *Reaction.*—Consists in operating the switch, and thus opening the direct connection from the battery to contact A_1,



SCHEMATIC I.

M. H. 12.29.30

the short-circuit is opened, the relay winding receives again current, operates its springs and thus breaks the contacts A_1, B_1 and A_2, B_2, thus breaking the marker circuit and re-establishing the high resistance stage in the main circuit. The relay is again being operated and while the current is going through the bulb it is too weak to light it.

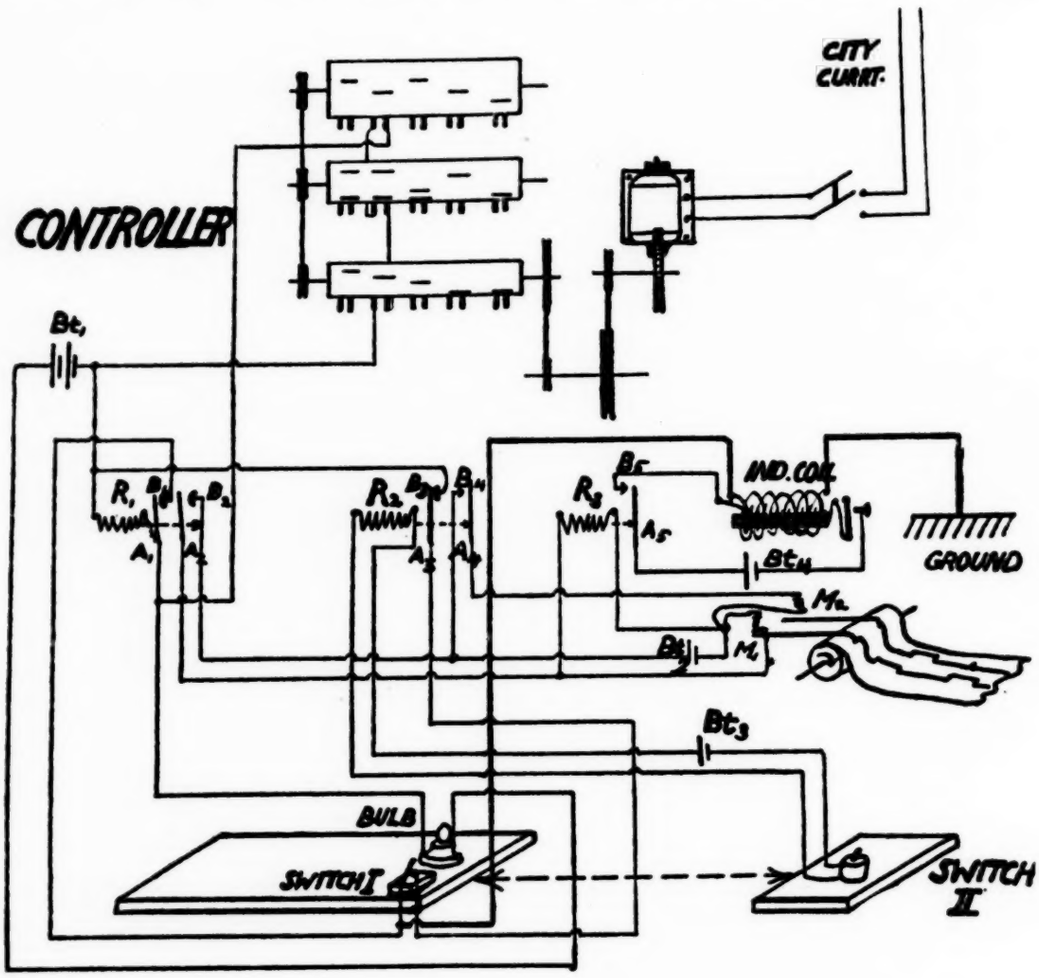
4. This again corresponds to the *rest stage*, *No. 1*.

Schematic II

Case of Signal 'A'

1. *Rest Period.*—A very weak current flows through the system Bt_1–R_1–Bulb–Bt_1; this current is sufficient to keep R_1 operating. The springs of the controller establish a direct connection between Bt_1 and the outgoing pole of the relay R_1, thus short-circuiting R_1 and causing it to release its springs. This initiates

2. *Operation.*—Circuit Bt_1–Bulb–A_1–B_1–Switch I–A_3–B_3–Bt_1 is closed. This is a permanent circuit which can only be



SCHEMATIC II.

M.H. 12.29.30

interrupted by opening either switch I or A_3–B_3 (by operating the relay R_2). R_1 causes A_2–B_2 to be closed, thus operating the marker M_1 (Bt_2–B_2–A_2–M_1–Bt_2) and the relay R_3. This relay in turn, closes the circuit Bt_4–B_5–A_5–Induction coil–Bt_4. This operated the induction coil, thus putting the switch I under tension.

3. *Reaction* is possible in *two ways:*

(*a*) If the switch I is pulled, although it is electrically charged, the principal circuit is interrupted and causes essentially the relay R_1 to receive current again, thus rendering the opening of the circuit permanent through separation of A_1 and B_1.

(*b*) The switch II may be operated; this closes the circuit Bt_3–Switch II–R_2–Bt_3. R_2 operates the springs A_3 and B_3, and interrupts thus the principal circuit. A marker M_2 is operated at the same time, as contact B_4 and A_4 is made (Bt_2–A_4–B_4–M_2–Bt_2).

As soon as R_1 resumes the active position, A_2 and B_2 are opened, both the marker M_1 and the relay R_3 are released; thus breaking the contacts A_5 and B_5, which stops the operation of the induction coil.

Both (*a*) or (*b*) lead up to the starting point (1) again.

(Manuscript received February 28, 1931)

APPARATUS FOR MEASURING ACTIVITY

BY THOS. H. HOWELLS

University of Colorado

One of the most fundamental characteristics of life and personality is activity. Measurement of the spacial and temporal characteristics of the movements of the body, taken as a whole, as well as of the component activities of the individual parts, constitutes one of the persistent problems in human and animal psychology. The various hypotheses concerning causes, rates, levels, cycles and patterns of activity cannot be tested without exact measurement of fractional and total activities. Neither can inductive reasoning occur in the absence of specific data. The purpose of this writing is to describe some means of obtaining more exact measurements of bodily movement.

Measurements of activity have generally been procured through attaching strings, levers, rheostats, or pneumatic devices either to the body of the *S*, or to the movable cage, platform or bed which supports the *S*. There are a number of deficiencies inherent in these systems among which are the following:

1. They are cumbersome and difficult to manage.

2. The deflection of the recording stylus is frequently not proportional to the amount of activity. Deflection may result from maintenance by the *S* of a changed static position.

3. The reference or zero line is constantly shifting in the typical activity record.

4. Summation measures of component or total activities are not provided except through calculations based upon measurements of oscillographs.

It is believed that the apparatus described below quite largely overcomes these difficulties. Two different types of activity cages for animals are first presented, followed by

descriptions of an electrical activity meter and a hydraulic activity meter.

Figure 1 shows a cubical cage with glass sides which is suspended by springs connected to the eight corners of the box. Figure 2 shows a rotary cage with a screen floor and

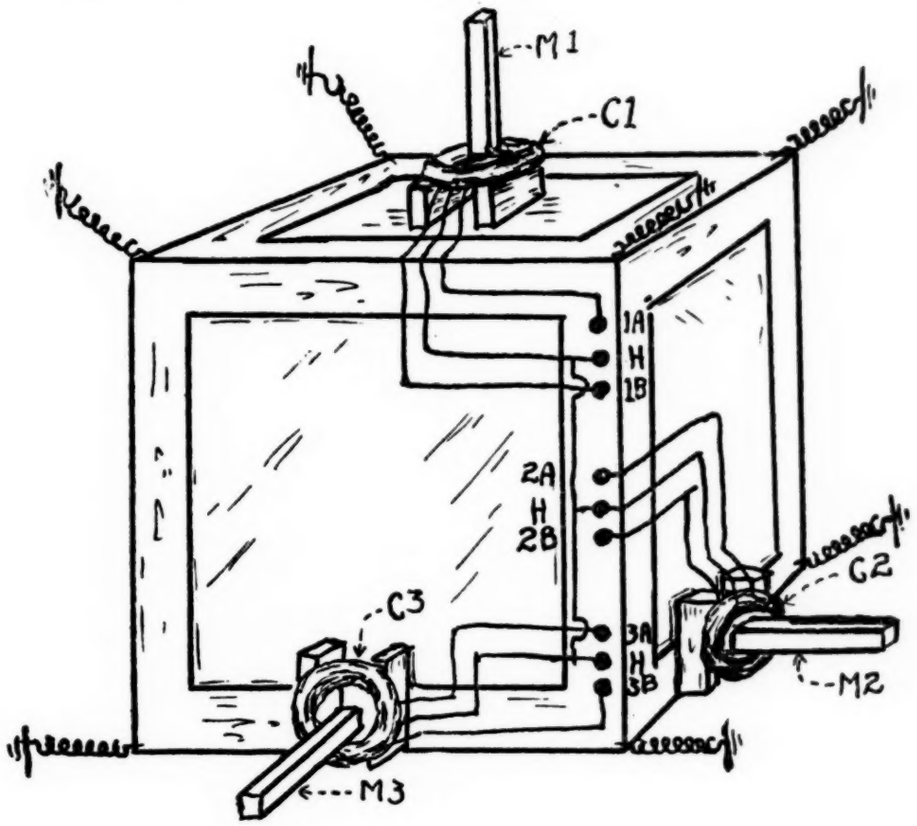


FIG. 1. Spring mounted activity cage.

which is mounted in a frame so that it is free to rotate. The frame is flexibly suspended. It is desirable to damp the movements of both cages by connecting them to movable vanes which are immersed in water.

When these cages are used the record of the movement of the animal is obtained indirectly from a record of the resultant motion of the cage. Attached to the cage of Fig. 1 are three coils of 1000 turns each of No. 40 insulated wire, hereafter referred to as pick-up coils. Mounted separately on the supporting frame are three associated, permanent bar magnets. One end of each magnet protrudes slightly through the center of the corresponding coil. The parts are so ar-

ranged that motion of the cage moves the coils with reference to the magnets and induces electrical potentials at the terminals of the wire. The coils are mounted far enough away from the cage so that movement of the cage will not cause the magnets to jam against it, and the inside diameter of the coils is great enough so that movement at right angles to the coils is possible. The coils are mounted on three

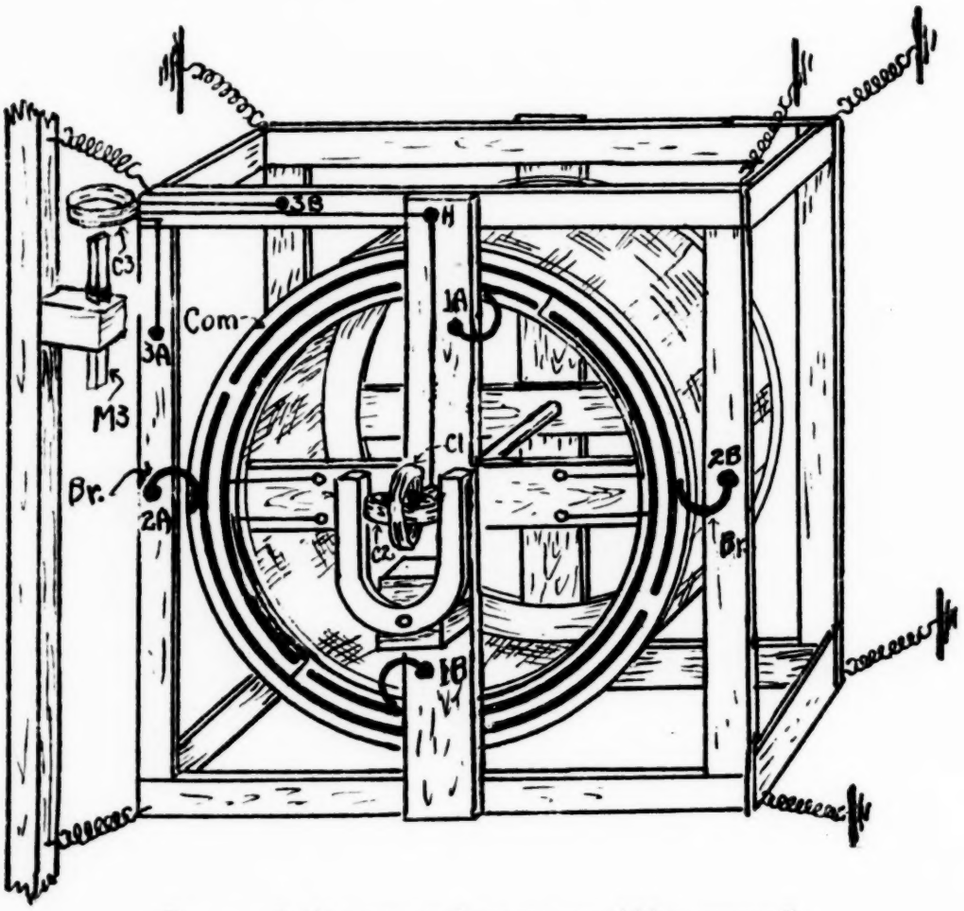


FIG. 2. Spring mounted rotary cage (sides removed).

faces of the cage so that movement in any one of the three dimensions must affect some one of them. These coils, together with the corresponding magnets, are shown in diagram at the top of Fig. 3. The terminals of these coils are indicated in all three figures by the black dots and are correspondingly numbered. It will be noted that each coil has a center tap.

The purpose of the apparatus which is diagrammed in Fig. 3 is to amplify, coördinate and measure the potentials

induced in the pick-up coils. These potentials are, of course, proportional to the movement of the cage. Each of the terminals of each coil is connected to the grid of one of a pair of radio vacuum tubes of the 226 type. The plate current through these tubes has been reduced to a minimum because of the influence of the *C* battery, *CB*. Whenever movement of any one of the pick-up coils occurs a positive

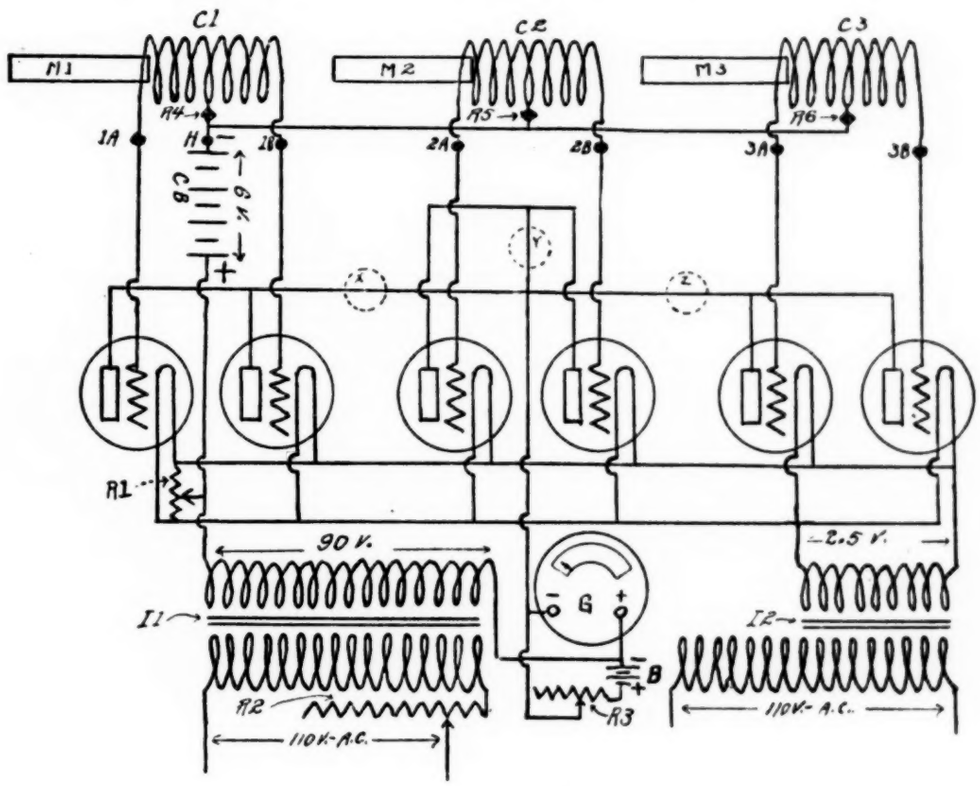


FIG. 3. Activity meter (electrical).

potential is induced on one or the other of its terminals, and on the grid of the connected tube. There is a consequent increase in the plate current flowing through that tube, which increase is registered on the galvanometer, *G*. On the other hand additional induced negative potentials on the grids of these tubes do not appreciably diminish the plate current flowing in such tubes. This is due to the fact that the tube is operated near the lower end of its characteristic curve, or in other words that the negative potential of the *C* battery has already reduced the plate current to a minimum. The result is that movement of any coil in any direction,

or any combination of such movements, tends always to move the needle of the galvanometer in the same direction and to increase its deflection. The meter therefore gives a summation measure of to and fro movements in three different directions.

*I*1, in Fig. 4, is an induction coil with the interrupter shorted and with a 100,000 ohm variable rheostat, *R*2,

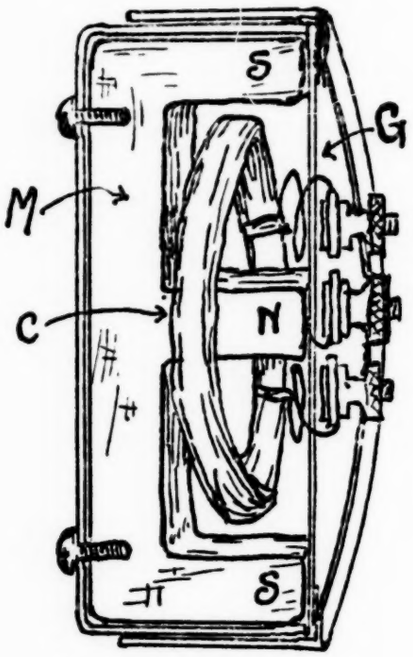


FIG. 4. Activity pick-up.

connected in series with the primary. The rheostat is adjusted until the proper output voltage is obtained.[1] *I*2 is any transformer that can supply the proper potential and current for the filaments of the tubes. An adjustable toy transformer serves very well. *R*1 is a 20 ohm potentiometer. *G* is a sensitive galvanometer of any sort, preferably of the mirror type if it is to be used for recording. The coil of the galvanometer should be wound to a resistance of 1000 ohms or over. If screen-grid tubes are used the resistance of the meter should be 10,000 ohms or over but the sensitivity will be multiplied several times. Connected in series across *G* is a 100 v. '*B*' battery and a rheostat (Resistograd, range

[1] A 90 v. *B* battery may be substituted for *I*2 if desired. The use of the battery will eliminate any tendency toward vibration of the elements of the galvanometer.

200 ohms to 30 megohms). This rheostat should be adjusted so as to neutralize the effect of the initial plate current and bring the meter to zero reading.

Additional meters can be connected at *X*, *Y* and *Z* in case separate measures of the component movements are desired. The relative influence of each of these components on the summation record can be controlled through adjustment of the resistances, *R*4, *R*5, *R*6, etc. Any number of pick-up units can be added if it is desired to include in the composite more than the three activity factors. The general plan of the diagram should be followed.

Use of the inductor pick-up units in connection with a cage, as in Fig. 1, is only typical. Figure 2 shows two coils mounted so as to turn with the rotary cage and induce potentials through interaction with the horseshoe magnet. A third coil and magnet are situated on one corner of the frame. Pick-up units of the type illustrated in Fig. 4 can be substituted in situations in which it is difficult to use the other type of unit. This unit resembles a watch case and can be mounted on any part of the body of a human or animal subject. It may also serve to measure the vibration of apparatus or machinery. The coil, *C*, in this unit, is flexibly mounted so that whenever movement occurs it will vibrate and generate an EMF through interaction with the magnet, *M*. It is provided with a glass front for the sake of easy inspection.

A similar apparatus can be used for the purpose of obtaining curves of sound intensity. In this case the coil and magnet units should be replaced by microphones, preferably of the inductor type. The same amplifier would be used. In case greater amplification were needed the tubes shown could be replaced by tubes of the screen-grid type.

Figure 5 shows a simple hydraulic type of activity meter. While it is not as sensitive or as flexible as the electrical meter, previously described, it does have most of the other advantages mentioned and is perfectly satisfactory for some purposes.

The level of the water in the trough, *T*, is held constant

because of flow through a siphon tube to a stand-pipe, or reservoir containing a large volume of water at the same level as in the trough. Valves are shown to prevent the flow of water back through the siphon.[2] The drive-rod, *D*, is connected to the subject or to the cage, and the movements

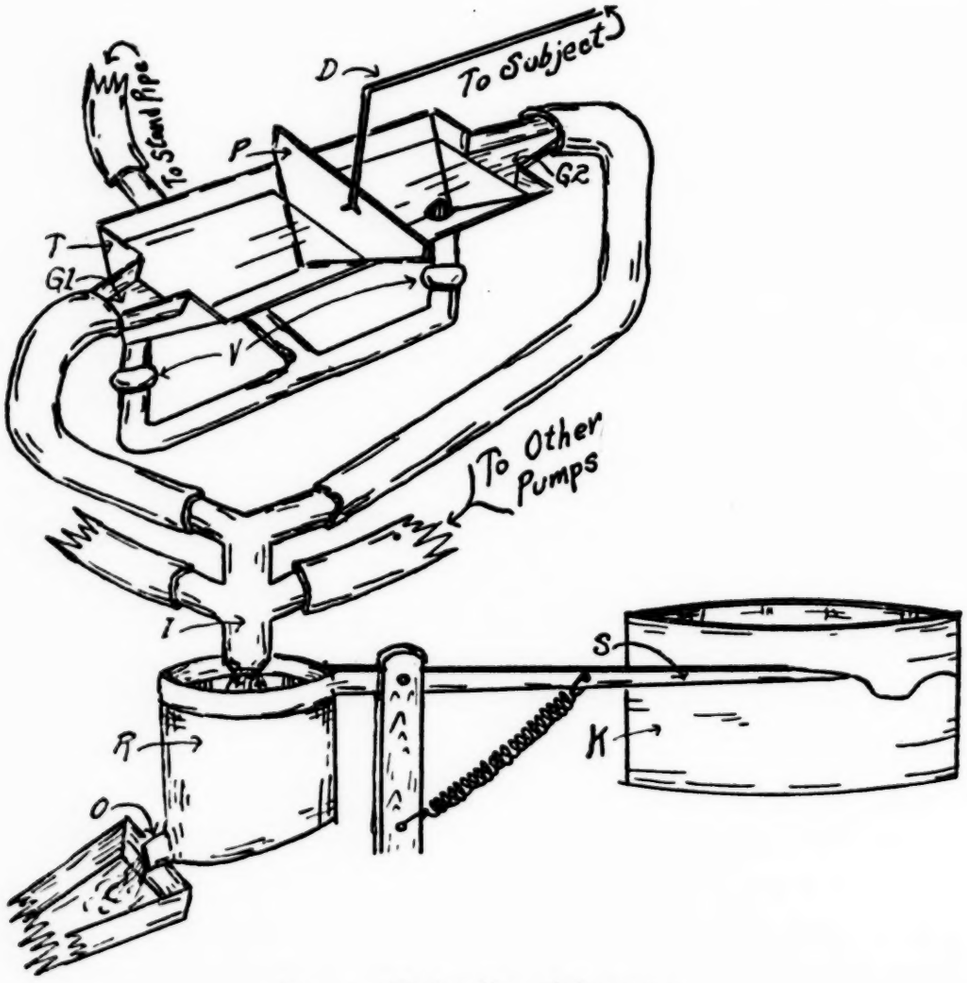


FIG. 5. Hydraulic activity meter.

push the plunger, *P*, back and forth, resulting in an overflow of water through the gates, *G*1 and *G*2. This water flows down through the rubber tubes into the receptacle, *R*1, which causes a deflection of the stylus, *S*. The amount of water displaced is proportional to the movement of the plunger and the connected subject. The latter is therefore registered directly, in terms of deflection of the stylus from the zero line, on the kymograph drum, *K*. The outlet, *O*, from *R*, must

[2] In practice these valves have been found to be somewhat unnecessary.

be large enough so as to retain the water long enough to produce a deflection of the stylus, and yet not so long as to cause periods of activity to overlap each other on the record.

With this apparatus there is an inevitable period of delay between occurrence of the activity and the time at which it is recorded. This lag in the record is uniform, however, and can be compensated for. It can be considerably lessened by coating the inside of the tubes with paraffin or vaseline, thus largely eliminating adhesive friction. As indicated in Fig. 5 the flow of water from other similar trough-units can be led into the same receptacle and join with the first in producing a summation record on the drum.

It will be noted that both of the described types of apparatus have the following advantages: (1) the line of reference does not change, (2) deflection of the record from this line is produced and maintained only by activity of the *S*, and (3) partial and summation records can be obtained of various movements.

(Manuscript received January 24, 1931)